The Dissertation Committee for Pranav Madadi
certifies that this is the approved version of the following dissertation:

Performance Analysis of Mobile Users in Poisson Wireless Networks

Committee:

François Baccelli, Co-supervisor

Gustavo de Veciana, Co-supervisor

Jeffrey Andrews

John Hasenbein

Sanjay Shakkottai

Performance Analysis of Mobile Users in Poisson Wireless Networks

by

Pranav Madadi,

DISSERTATION

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

DOCTOR OF PHILOSOPHY

THE UNIVERSITY OF TEXAS AT AUSTIN

August 2019

Dedicated to my parents and brother.

Acknowledgments

The four years of the Ph.D. program at UT Austin was an incredible journey in my life, and I would like to express my gratitude to many people for their help and supports during this time.

First, I would like to extend my sincere gratitude to my advisors Prof. François Baccelli and Prof. Gustavo de Veciana. Their approach towards new research topics, immense knowledge on a multitude of topics, work ethic, general enthusiasm towards research have been very inspiring. They have been very patient with me and great support at times of difficulty with research. They have guided me not just on topics of research but also on deciding on my career goals, honing my presentation and writing skills. I have developed a deep mathematical understanding from various courses I have taken with Prof. Baccelli, who is an excellent teacher. The mere hard work and enthusiasm showed by Prof. Gustavo in teaching always amazed me and it was an enriching experience to be a teaching assistant for his course. He is also an excellent teacher who thrives on providing unique insights and I had a great time taking his course on queueing theory. They played a huge part in shaping me as an independent researcher over the past five years and I am very grateful for the same.

I would like to thank professors in my dissertation committee, Prof.

Sanjay Shakkottai, Prof, Jeffrey Andrews, and Prof John Hasenbein for providing helpful feedback for my research and dissertation. I have developed a great understanding of the concepts of probability from the course I had with Prof. Sanjay Shakkottai. His excitement about various topics is contagious and inspiring. I am grateful for various insightful discussions with fellow Ph.D. students in my lab. I would like to extend my gratitude to Virag, Arjun, Abhishek, Rajat, Shalmali, Avradeep, Junse Lee and Surya. I would like to thank Melanie Gulick, Karen Little and Melody Singleton for helping me with a plethora of issues.

I have grown not just as a researcher but also as a human being and have had a great time in Austin for the past five years, thanks to an amazing set of people. Vatsal Shah has been a constant in every endeavor of my Ph.D. life and has helped me navigate through troubled times. The lessons I have learned through numerous discussions we had shaped me to be a better person and for that, I am forever grateful. I would like to thank Priyanka Khante for always lending a listening ear to my pettiest grievances without passing any judgment. Vikas Reddy has played an instrumental role in my seamless transition to life in Austin and I will always cherish numerous memories we shared as roommates. Nitin Prasad, Venktesh Pandey have been a constant source of encouragement in embracing a fit and active lifestyle. Thanks to Bharath Bangalore for organizing all the lunches on campus. Without Soumya, Preeti, Prashu, Esha, Roshan, Priya, and Luca I would not have lived such a colorful life filled with parties and fun. The memories we created on the

many many road-trips, the long night-outs that often ended with lively debates and discussions, are something that not only helped me navigate through my Ph.D. but are something that will stay with me for the rest of my life and I am immensely grateful to all my friends for being an integral part of that.

My professors and friends at IIT Hyderabad provided never-ending support and encouragement in my decision to pursue a Ph.D. I am especially grateful towards Prof. Kiran Kuchi, my B.Tech project advisor, for sparking my interest in mobile communications. I would like to thanks Prof. Sumohana and Prof. Soumya Jana for guiding through my Ph.D. applications and being great teachers. I have been blessed to be around an amazing set of friends who have been very supportive. Ameya Patil always pushed me to strive for excellence, helped me make up my mind to pursue my Ph.D. and was always there for me when I needed him. Nayan Panaganti, taught me how to be an amazing friend by being one. I would like to thank Deepak, Bharat, Krishnan, Eswaran, Karthik, Manikanth, Vishnu and Nachiket for all the memories.

I will be eternally grateful to my parents Ram Mohan and Padma; they not only supported me in all my endeavors but also were the ultimate pillar of strength through all my highs and lows with their unwavering confidence and belief in all my decisions. I feel extremely lucky to have parents who motivated me to pursue my dreams by being there at every step. I would like to thank my brother Prakyath, who always inspired me with his disciplined life. Alekya, Meghana, Monica, Anvitha, Aishwarya, Soujanya, Saroj, Tharuni, Srinivas thank you all for filling my life with joy and support. I would like to thank

my brother Ashwin for inspiring me to pursue IIT.

Performance Analysis of Mobile Users in Poisson Wireless Networks

Publication No. _____

Pranav Madadi, Ph.D.
The University of Texas at Austin, 2019

Supervisors: François Baccelli
Gustavo de Veciana

Stochastic geometry is a widely accepted mathematical tool used to analyze cellular networks, where the location of base stations are modeled by spatial point processes. It is used to derive closed-form or semi-closed-form expressions for the SINR or for the functions of the SINR which determine various network performance metrics such as coverage probability, "edge" capacity, 90% quantile rate, spectral efficiency, and connectivity without resorting to complicated simulation methods.

Predominantly, it is used in deriving marginal distributions of SINR by considering a typical user assumed to be located anywhere on the plane. Models beyond the typical user approach have been proposed with the aim of analyzing QoS metrics of a population of users, and not just a single user. Most of which include considering networks at certain times by representing

instances or snapshots of active users as realizations of spatial (usually Poisson) processes or users occurring at random locations that last for some random duration. Analyzing the performance of a typical mobile user on the move or that of a population of such mobile users is complicated since it requires studying not just the marginal but the spatial stochastic fields associated with wireless networks.

In this thesis, we model and analyze the fields associated with wireless networks where the locations of base stations are distributed according to a homogeneous Poisson point process. We focus on characterizing the level crossings, extremes, and variability of the Shannon rate fields in noise limited (SNR based) environment by establishing a connection to queueing theory. In interference limited (SIR based) environments, we rely on the theory of Gaussian random fields which arise as natural limits of standardized interference under densification. Using this, we characterize the spatial correlations, and variability of the Shannon rate fields in the limiting regime. We leverage the spatial characterization of the fields to study the temporal variations and various Quality of Service (QoS) metrics seen by the users on the move. The quantification of such metrics as a function of a small number of network parameters, e.g., the density of base stations, path loss, should allow network operators to appropriately tune the density of the base stations to meet the demands of mobile users for a required performance level.

In noise limited environments, we study the performance of mobile users in dense networks by incorporating the cost of handovers along with the

temporal variability in the Shannon rate. We study the tradeoff between the cost of handover and the Shannon rate by proposing a new class of association policies. Associating with a base station that is farthest in the known users' direction of motion leads to fewer handovers but may lead to a decrease in the rate. Thus, we attribute a local association region to the mobile user to restrict the greediness in the association, which also models the constraint on the available information about the locations of the base stations. We propose a class of greedy association policies and once again leverage stochastic geometry to characterize the performance of such policies. We then optimize the shape and size of the association region by establishing a connection to the theory of Markov processes and compare the performance of this policy to traditional association policies.

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

The high rate of technological advancement in the wireless industry has resulted in a substantial increase in the use of mobile devices such as smartphones and tablets. Indeed cellular traffic generated by mobile devices has grown 18 fold in the past 5 years and by 63% in 2016 alone [1]. This growth is driven by increased smartphone subscriptions and increase in the average data volume per subscription, fueled primarily by more viewing of video content.

The convenience of mobile devices (smartphones) resulted in their intensive use during commute time. Approximately 20-30% of cellular data is generated during commute hours by wireless devices on the move [2]. Applications like video/audio streaming, which accounted for 60% of total mobile data traffic in 2016 [1], geolocation services/maps, and social media are extensively used by users during commute periods. In the future, with the increase in the use of public transportation and the emergence of self-driving cars, this volume could grow substantially.

Mobile users expect excellent coverage with high speeds for their applications to run smoothly anytime/anywhere. Thus, it is a challenge for service

providers to deliver high capacity to a large volume of users especially the ones on the move. To meet with such increasingly stringent performance requirements of users along with the increase in traffic volume, various technologies are being explored for 5G networks. Although millimeter wave and massive MIMO have been proposed to improve spectral efficiency and exploit wider bandwidths, it is expected that Ultra Dense Networks (UDNs) will be the key tool towards boosting capacity and coverage. Thus, it is of utmost importance to focus on improving the performance seen by users on the move in such environments, i.e., ultra dense networks.

Given that the quality of experience seen by users on the move depends on many factors other than the link capacity, we analyze two main contributors. The first is associated with temporal capacity variations mobile users would experience. This is particularly the case for applications that are delay sensitive, e.g., video/audio streaming and for real-time services, e.g., navigation/augmented reality, which may not be able to smooth substantial capacity variations through buffering. The main challenge in understanding such temporal variations is in modeling the continuous time stochastic processes associated with wireless networks such as the Shannon rate and interference fields.

The second contributor is linked to the cost of frequent handovers. The high frequency of handovers resulting from mobility in dense networks can be detrimental to user performance. Unfortunately, the negative impact of handover overheads in dense networks when serving mobile users is often ig-

nored [3]. A high handover rate results in high overhead for both applications and infrastructure which is, of course, exacerbated in dense networks. In addition to the handover signaling overheads, the handover procedure interrupts the data flow to the users due to link termination with the serving base station and link establishment with the target base station. A recent study [4] showed that handovers generally cause short-term disruptions in various applications. Similarly, simulation studies [5] have shown that handovers degrade the performance of real-time applications such as VoIP. Thus, the increased rate of handovers diminishes the benefits of base station densification.

In this thesis, we focus on analyzing the performance seen by mobile users by studying the above-mentioned contributors. The first challenge then is to model the wireless network, the capacity variations, and the handovers seen by mobile users. Stochastic geometry is a widely accepted mathematical tool to model and analyze cellular networks which enables performance characterization in terms of the base station intensity, λ as well as other physical layer parameters (see [6] for a survey). In the sequel, we will consider the situations where base station locations are distributed according a homogeneous Poisson point process.

We consider the simplest meaningful model for the tagged mobile user, i.e., moving along a straight line at a fixed velocity and characterize the temporal stochastic processes as a cross section of the associated random fields in various environments. Later, we propose mobility-driven association policies and study the interplay between handover cost and capacity gain as a function

of the base station intensity. Thus, the aim of this thesis will be to develop a better understanding of the design challenges for dense wireless networks, supporting large numbers of mobile users.

1.1 Summary of Accomplished Work

This thesis is organized as follows. The first two chapters focus on developing a framework to study the variations in the Shannon rate fields of Poisson wireless networks in various environments such as noise-limited and interference-limited dense networks. We consider various practical communications engineering questions where the developed framework can be directly used such as adaptive coding and modulation for mobile users, the problem of backhaul provisioning in ultra-dense networks, stored-video streaming, and WiFi offloading. The last chapter focuses on mobility-driven association policies and validating their promise and effectiveness.

1.1.1 Chapter 2: Shannon Rate Process in Noise Limited Environments

In this chapter, we focus our attention on a tagged user in the following context: (a) the access points/base stations are a realization of a Poisson point process; (b) other users sharing the network also form a Poisson point process; (c) in this initial work, we assume users associate with the closest base station; (d) resources(time) is shared equally among users sharing the same base station, and (e) a standard distance based path loss is used with neither

shadowing nor fading. Our goal is to characterize the shared rate process seen by the tagged user by studying the variations in the SNR and variations in the number of users sharing the infrastructure.

We begin with a characterization of the SNR process. Given a SNR threshold, we provide a complete characterization of the on/off level crossing process as an alternating-renewal process. Equivalently, this characterizes the duration for coverage/outage events seen by the mobile. We then provide asymptotics for the likelihood of high and low SNR, and show that, after appropriate rescaling, the time intervals between up crossings are exponential, which provides a new illustration of the rarity hence exponentiality principle [7], [8].

Next, we consider the sharing number process, i.e., the number of other users which meet an SNR threshold, and share infrastructure nodes with the tagged user. The contrast between the underlying mechanism leading to the occurrence of rare events associated with high and low shared rates is highlighted. We also provide an asymptotic characterization for the re-scaled inter-arrival times associated with up-crossings.

Our results are then leveraged to evaluate the QoE experienced by mobile users in two concrete scenarios, which gives a system-level view on the performance that mobile users would see.

The first scenario considered is the delivery of streaming video to mobile users which are able to pre-buffer future video frames to prevent rebuffering.

The primary question addressed is a characterization of the maximal spatial density of users that can be supported while ensuring that, in the long term, no rebuffering is required. Our second application considers the distribution for the delays experienced by a mobile user attempting to download a large file, which gives a system-level view on the performance that mobile users would see when downloading large files opportunistically via some WiFi infrastructure.

We briefly discuss some important extensions of our results to the case of heterogeneous wireless infrastructures and the case where the locations of mobile users follow a Cox process associated with a road network. The results show how the use of heterogeneous technologies will impact the temporal variations in the mobile users SNR process. The results also show how a mobile user on a roadway is likely to see poorer performance than a static or pedestrian user.

Results based on discrete event simulation results are also presented. These are used to assess the robustness of this model to perturbations that were not yet taken into account in the analysis. For example the relative impact of variability associated with the changing geometry (proximity of base stations) seen by a mobile versus that associated with channel variability due to channel fades. Further, the robustness of our model to changes in the direction of the motion of the mobile user i.e., versus straight line motion, is studied. We complete the chapter by discussing proposed future work.

1.1.2 Chapter 3: Shannon Rate Process in Interference Limited Environments

In this chapter, we study the spatial characterization of the Shannon rate field of a Poisson network under densification. To begin with, a scaling limit for interference field under bounded, non-negative integrable path loss functions is established. This is then used to approximate the interference field in dense networks by a stationary Gaussian field which captures the underlying spatial variations.

Since the interference field primarily depends on the distribution of the base station locations and the path loss, we classify various existing path loss models in the literature such as the dual slope models ([9], [10], [11]) and observe the impact they have on the various sample path properties of the field, e.g., continuity and differentiability.

By transforming or taking functionals of the interference field, one can study various additional properties of such systems. In particular, with an appropriate bandwidth scaling, we model the Shannon rate field as a stationary Gaussian field. In turn, relevant functions of the Shannon rate field enable the study of variability in the spatial average rate, and backhaul capacity requirements for ultra dense wireless networks. We explore how various sources of variability impact backhaul dimensioning.

We then study how the spatial variability of the interference and Shannon rate field impact the temporal performance variations a mobile user would see. For certain path loss models, the interference process is nowhere differ-

entiable, which implies that the mobile experiences high fluctuations in rate making it difficult to implement adaptive modulation and coding techniques. To better understand these fluctuations, we quantify their variation via Hölder exponents [12] and leverage these to bound the time scales over which adaptive modulation and coding could be performed. Finally we complete our work by discussing a characterization of the level crossing characteristics, e.g, interference, and or rate properties of such fields as experienced by a mobile users is provided. We complete this chapter by discussing further proposed work in this area.

1.1.3 Chapter 4: Mobility-driven Association Policies

In this chapter, we study the problem of determining optimal base station association policies which tradeoff throughput with handover penalties. We consider the noise-limited environment where the base stations are distributed according to a Poisson point process. We model the cost of a handover by considering the loss in throughput associated with handover delays. We also introduce an additional fixed loss in the data per handover to model the packet loss and signaling overhead.

In our first attempt, we consider a setting where the mobile user's trajectory is known as well as the direction of motion and the location of all base stations in space. In this setting, we consider the total volume of data delivered to the mobile user accounting for handover overheads during the mobiles' trajectory as our performance metric. We show that the problem of

optimizing this metric is challenging and, under some simplifying assumptions, develop a dynamic programming problem that reduces the above optimization problem to finding the maximum weight path in a directed acyclic graph.

Given users are often constrained with the information about the base station, we consider the setting with high handover costs and constrained information about the base station locations. For this case, we propose a greedy approach where the user associates with the base station that is farthest in the direction of its motion, i.e., greedily pick the base station that results in the maximum connection time. The constraint on the knowledge of base stations also limits the greediness. We model the information available to the tagged user with respect to a geometrical region, i.e., all the base stations within a specific region centered at the user. Here, we consider the time average of the throughput as the performance metric and optimize it with respect to the parameter associated with the geometric region.

Under our proposed greedy association policy, using the properties of Poisson point process, we show that the evolution of the mobile user's association base station is Markovian over time. More precisely, we establish a connection between the random geometry of base stations and the theory of Markov processes with respect to time. We then study various properties of this continuous state space Markov process such as its irreducibility, aperiodicity, and characterize its stationary distribution for one specific class of association policy.

Using various results on Markov processes such as the cycle formula [13],

we evaluate the average throughput seen by the mobile user in the closed form. We show that there exists an optimal size of a given geometric region maximizing average throughput and evaluate it. We study the trade-off between the farthest distance and the height of the handover support set with respect to the performance for the various intensity of base stations. With the help of simulation, we then compare its performance with the optimal performance of the previous dynamic programming solution and closest base station association policy.

Chapter 2

Shannon Rate Process in Noise Limited Environments

2.1 Introduction and Related Work

The primary aim is to model and study the temporal capacity variations experienced by wireless users moving through space. Our focus on temporal variability should be contrasted with the extensive work characterizing spatial variability as seen by randomly located users. The temporal variations are a result of not just the variability in channel strength such as fading, shadowing and path loss but also from the geometric variations in the spatial and environmental relationships (associations) to the infrastructure, the number of users sharing the network and variations in interference. Although space and time may be related through averages (when ergodicity holds), the temporal characteristics of the stochastic processes modeling a mobiles capacity vari-

Portions of this chapter has been published as: Madadi, Pranav, Francis Baccelli, and Gustavo de Veciana. "Shared rate process for mobile users in poisson networks and applications." IEEE Transactions on Information Theory 64, no. 3 (2017): 2121-2141.
Co-authors have participated extensively in model formulation and research methods, and have contributed in reviewing the final manuscript.

ations are relatively unexplored beyond correlation analysis and are of great practical interest.

Key Questions. Two primary sources of temporal variation in a mobiles capacity are changes in the SNR which are associated with changes in the users geometric relationship to the infrastructure, and the sharing number, i.e., the number of other mobiles sharing the resource. Our goal in this chapter is to study the relative impacts these have on mobile users QoE. In this setting several basic questions arise:

1. ***Characterization of the SNR process.*** As a mobile user moves through a wireless network infras- tructure associating with the closest node, its SNR process and thus associated maximal rate (i.e., without accounting for sharing) will experience peaks and valleys. What is the intensity of peaks and valleys? Does it match the rate of cell boundary crossings? Given an SNR threshold, can one characterize the temporal characteristics of the on/off level crossing process associated with being above and below the threshold, i.e., the coverage and outage durations?

2. ***Characterization of the sharing number process.*** Assuming a population of other users sharing the network, what are the characteristics of the sharing number process seen by a mobile? If the network is shared by heterogeneous users, i.e., static, pedestrian, and users on public transport and/or a road system, how will this bias what a mobile user sees?

3. ***Smoothness of the effective rate process.*** How smooth is the bit rate obtained by the mobile seen as a function of time? How often does this rate incur discontinuities, trend changes, large jumps or other phenomena negatively impacting real time applications?

4. ***Characterization of rare events.*** Conditional on a rare event, i.e., very poor or very good user rate, what is the relative contribution of the users location vs network congestion? Can one characterize the time scales for rare event occurrences?

5. ***Applications to QoE.*** What are the implications of the temporal capacity variations mobiles see on their application-level QoE and system-level performance, e.g., acceptable density of video streaming users it can support, or download delays of large files?

To the best of our knowledge the analysis of the SNR and shared rate processes developed in this chapter for Poisson wireless infrastructure is new and provides a first order answer to these basic questions. While our models are simplified they should be viewed as a first, and necessary, step towards characterizing temporal variability in more complex random structures, e.g., SINR variations which, in this chapter, are only explored via simulation for comparison to SNR process characteristics. Generalizations to SINR processes would be desirable, particularly for systems operating in the interference limited regime. Such extensions may be tractable based on results known for shot noise fields (see [14], Volume 1), yet are beyond the scope of this chapter.

Related work. There is a rich literature on modeling spatial capacity variability in wireless services experienced by a randomly located user. Of particular relevance is that based on stochastic geometry, which captures the effect of the variability in base station locations, as well as the variability in the environment through shadowing, and in the channel through fading, see e.g., [14] for a survey.

There is also significant related work on Delay-Tolerant Networks (DTN). This literature considers mobile nodes, where the contact duration and the inter-contact times are defined and empirically measured from real traces [15], [16] as well as through underlying mobility models [17]. However, this work is geared at studying opportunistic ad-hoc communication networks where the main question is the end to end delay of a tagged packet, under the assumption that nodes are moving independently. Our focus is on studying the continuous-parameter stochastic process experienced by a tagged mobile user traversing a static pattern of nodes e.g., base stations, modeling the wireless infrastructure.

There certainly is a lot of interest in studying how to design networks to better address the needs of mobile users or to leverage user mobility for the offloading/onloading traffic. For example [18] studies how user mobility patterns and user perceived QoE might drive the selection of macro-cell upgrades. The work in [19] examines the effectiveness of algorithms for optimizing offloading to a set of spatially distributed WiFi APs. The work in [20] evaluates how proactive knowledge of capacity variations could be used in designing new models for video delivery. These works exemplify applications and engineer-

ing problems which depend critically on the temporal capacity variability that mobile users would experience, but do not directly address the characteristics of such processes.

2.2 System Model

Consider an infrastructure based wireless network consisting of nodes e.g., base stations/WiFi hotspots, denoted through their locations on the Euclidean plane. The configuration of the nodes is assumed to be a realization of Poisson point process $\Phi = \{X_1, X_2, ..\}$ in $\mathbb{R}^2$ with intensity λ. We consider a tagged user moving at a fixed velocity v along a straight line starting from the origin at time $t = 0$. The mobile user shares the network with other (static) users which are spatially distributed according to another independent Poisson point process of intensity ξ.

All users associate with the closest infrastructure node. For each $X_i \in \Phi$ one can define a set of locations which are closer to X_i than any other point in $\Phi \setminus X_i$. This is a convex polygon known as the Voronoi cell associated with X_i [14]. The collection of such cells forms a tessellation of the plane called the Voronoi tessellation see e.g., Fig.2.1. Thus, all the users located within the Voronoi cell of node X_i associate with X_i.

Let $(X(t), t \geq 0)$ denote the random process, where $X(t) \in \Phi$ is the closest node to the mobile user at time t. Let $(L(t), t \geq 0)$ be the process denoting the distance from the mobile user to its closest node.

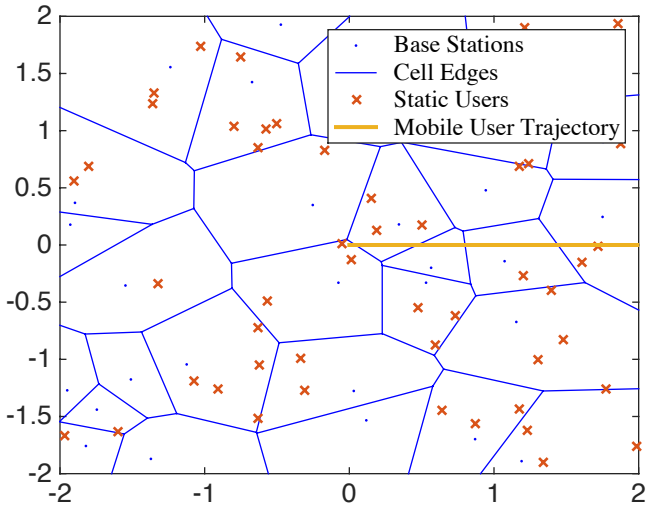


Figure 2.1: Mobile user motion in a sample of the Poisson cellular network.

We consider downlink transmissions and assume that all nodes transmit at a fixed power p. Based on the classical power law path loss model and in the absence of fading and shadowing, the *Signal-to-Noise Ratio* (SNR) *process* of the mobile user is

$$\text{SNR(t)} = \frac{pL(t)^{-\beta}}{w}, \quad \text{for} \quad \beta > 2, \quad t \geq 0, \tag{2.1}$$

where w denotes the noise power.

We assume that users are only *served* if their SNR exceeds a given threshold γ. It follows from (2.1) that a user is served if it is within distance $r_\gamma = (\frac{p}{w\gamma})^{1/\beta}$ from the closest node. We refer to r_γ as the *radius of coverage*. Note that a user may thus associate with a node but not be served.

The *Shannon rate process*, $(R^{(\gamma)}(t), t \geq 0)$, seen by the mobile user is

directly determined by the SNR process through the relation:

$$R^{(\gamma)}(t) = \begin{cases} a \log\left(1 + \mathrm{SNR}(t)\right) & \text{if } \ L(t) \leq r_\gamma, \\ 0 & \text{otherwise,} \end{cases} \tag{2.2}$$

where a is a constant depending on the available bandwidth.

We assume that each infrastructure node shares time equally among the users it serves. We define the *sharing number process* $(N^{(\gamma)}(t), t \geq 0)$, where $N^{(\gamma)}(t)$ is the number of static users sharing the infrastructure with the mobile user if it is served, and 0 if it is not served. In other words, if the mobile user is served by its closest node $X(t)$ at time t, it shares the resource with $N^{(\gamma)}(t)$ static users. The *sharing factor* $(F^{(\gamma)}(t), t \geq 0)$, is then defined by:

$$F^{(\gamma)}(t) = \frac{1}{1 + N^{(\gamma)}(t)}. \tag{2.3}$$

Finally, the *shared rate process* seen by the mobile user, $(S^{(\gamma)}(t), t \geq 0)$, is given by

$$S^{(\gamma)}(t) = R^{(\gamma)}(t) \times F^{(\gamma)}(t). \tag{2.4}$$

Our aim is to characterize this process. In the next two sections, we first study the two underlying processes namely: (1) the SNR process $(\mathrm{SNR}(t), t \geq 0)$ and its level sets, and (2) the sharing number process $(N^{(\gamma)}(t), t \geq 0)$.

In the sequel, for any stationary random process, e.g., $(S^{(\gamma)}(t), t \geq 0)$, we let $S^{(\gamma)}$ represent a random variable whose distribution is the stationary distribution of the associated process.

A summary of the key notation is provided in Table 2.1.

Symbol	Definition
Φ, λ	Poisson point process of nodes and its intensity
ξ	Intensity of users
p	Constant power transmitted by the nodes
$(X(t), t > 0)$	Random process denoting the closest node to the mobile
$(L(t), t > 0)$	Random process denoting the distance to the closest node
$(\mathrm{SNR}(t), t > 0)$	Signal-to-Noise ratio random process
$(C^{(\gamma)}(t), t > 0)$	Signal-to-Noise ratio level crossing process
$(R^{(\gamma)}(t), t > 0)$	Shannon rate random process
$(N^{(\gamma)}(t), t > 0)$	Random process denoting the number of users sharing
$(S^{(\gamma)}(t), t > 0)$	Shared rate random process
γ	Threshold on Signal-to-Noise ratio
r_γ	Radius of coverage for threshold γ

Table 2.1: Table of Notation.

2.3 Characterization of the Temporal Processes

2.3.1 SNR Process

This section is structured as follows: we start by considering the on-off coverage structure of the SNR process and then study in more detail the characteristics of its fluctuations. We then analyze the scale of the inter-occurrence times of certain rare events.

2.3.1.1 Analysis of the SNR Level Crossing Process

A first order question is whether the mobile is covered or not. To that end we define the SNR level crossing process as follows:

Definition 1. *Given an SNR threshold γ, the SNR level crossing process $(C^{(\gamma)}(t), t \geq 0)$ is defined as $C^{(\gamma)}(t) = \mathbf{1}(SNR(t) \geq \gamma)$ as shown in Fig 2.2.*

Clearly this is an on-off process, where the on and off periods correspond to coverage and outage periods respectively. The process alternates between "on" intervals of length $(B_n^{(\gamma)}, n \geq 1)$ and "off" intervals of length $(I_n^{(\gamma)}, n \geq 1)$ as depicted in Fig. 2.2. We also define the sequence of SNR up-crossing times $(T_n^{(\gamma)}, n \geq 1)$ for a given threshold γ. Let $V^{(\gamma)} \sim T_n^{(\gamma)} - T_{n-1}^{(\gamma)}$ be a random variable whose distribution is that associated with up-crossing inter arrivals as illustrated in Fig. 2.2.

In order to characterize the SNR level crossing process, we establish a connection between the time-varying geometry seen by the mobile user and an associated queueing process.

Let $D^{(\gamma)}(t)$ denote the closed disc of radius r_γ centered on the mobile user's location at time t. This closed disc follows the mobile user's motion along the straight line. Let $K^{(\gamma)}(t) = |\Phi \cap D^{(\gamma)}(t)|$ denote the number of infrastructure nodes in the disc.

The following theorem provides a simple characterization of $(K^{(\gamma)}(t), t \geq 0)$ which in turn will enable the study of the SNR level crossing process.

Theorem 1. *The process $(K^{(\gamma)}(t), t \geq 0)$ is same as that modeling the number of customers in an $M/GI/\infty$ queue with arrival rate $\lambda^{(\gamma)} = 2r_\gamma v\lambda$ and i.i.d. service times with density*

$$f_{W^{(\gamma)}}(s) = \begin{cases} \frac{v^2 s}{2r_\gamma\sqrt{4r_\gamma^2 - v^2 s^2}} & \text{for } s \in [0, 2r_\gamma/v], \\ 0 & \text{otherwise.} \end{cases} \tag{2.5}$$

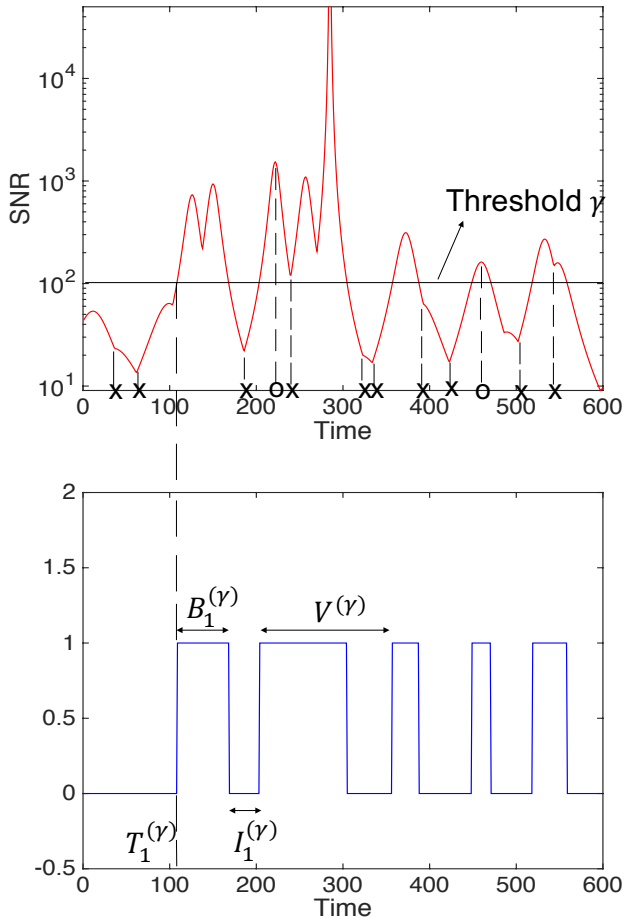


Figure 2.2: Level crossings of the SNR process as an on-off process (bottom) and the point process of its local maxima χ, denoted by "o" on the x axis, together with edge crossings, denoted by "$\times$" (top).

Proof. The entry of a node into the closed disc $D^{(\gamma)}(t)$ can be viewed as an arrival to the queue. The amount of time spent by the node in $D^{(\gamma)}(t)$ corresponds to its service time and thus the exit from $D^{(\gamma)}(t)$ to its departure from the queue.

We first show that the arrival process to the disc and thus the queue,

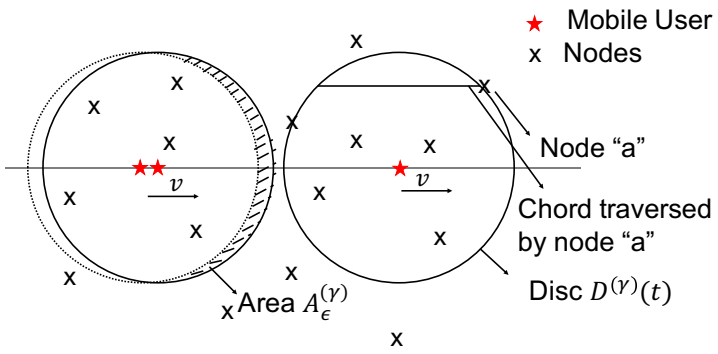


Figure 2.3: The disc around the mobile user with area $A_\epsilon^{(\gamma)}$ and the chord along which the node moves.

is Poisson. We begin by proving that the arrival process has independent and stationary increments. The probability that there is an arrival in the next ϵ seconds is the probability that there is a node in the set with area $A_\epsilon^{(\gamma)} = 2r_\gamma v\epsilon$ depicted in Fig. 2.3. Since nodes are distributed according to a homogeneous Poisson point process of intensity λ, the number of nodes in any closed set of area b follows a Poisson distribution with parameter λb. Thus, for any $\epsilon > 0$, the increments in the arrival process have the same distribution. Also, the number of nodes in any two disjoint closed sets are independent. Thus, the arrival process has independent stationary Poisson increments. For a small value of ϵ, the probability that there is a single arrival is $\lambda A_\epsilon^{(\gamma)} + o(\epsilon)$. Thus, the arrival process is Poisson with rate $2r_\gamma v\lambda$.

Every node that enters the moving closed ball stays in it for a time that depends on its entry locations and is proportional to the chord length as shown in Fig. 2.3. This corresponds to its service time in the queue.

Since the mobile moves at constant velocity, the distribution of the service times can be derived from the distribution of the chord lengths, see e.g., 2.3. Without loss of generality suppose the mobile moves along the x-axis. The y coordinate of a typical node entering the disc, $Y^{(\gamma)}$, is uniform on $[-r_\gamma, r_\gamma]$. The random variable $W^{(\gamma)}$ representing the typical service time can be represented as:

$$W^{(\gamma)} = \frac{2\sqrt{r_\gamma^2 - (Y^{(\gamma)})^2}}{v}.$$

The density of $W^{(\gamma)}$ is then given by (3.21), and has a mean $\mathbb{E}[W^{(\gamma)}] = \frac{\pi r_\gamma}{2v}$.

Thus, the process $(K^{(\gamma)}(t), t \geq 0)$ capturing the number of nodes in the moving disc follows the dynamics of the number of customers in an $M/GI/\infty$ queue with arrival rate $\lambda^{(\gamma)} = 2r_\gamma v\lambda$ and i.i.d service times following the distribution of $W^{(\gamma)}$. It follows that the stationary distribution for $K^{(\gamma)}(t)$ is Poisson with mean $\pi r_\gamma^2 \lambda$. □

Given this connection to an $M/GI/\infty$ queueing model, we can now characterize the SNR level crossing process as an alternating renewal process defined as follows:

Definition 2. *A process alternating between successive on and off intervals is an alternating renewal process if the sequences of on period $(B_n^{(\gamma)}, n \geq 1)$ and off period $(I_n^{(\gamma)}, n \geq 1)$ are independent sequences of i.i.d. non-negative random variables.*

Theorem 2. *For any $\gamma > 0$, the SNR level crossing process, $(C^{(\gamma)}(t), t \geq 0)$, is an alternating-renewal process. Further, its typical on period, $B^{(\gamma)}$, and*

off period, $I^{(\gamma)}$, are distributed as the busy and idle periods of an $M/GI/\infty$ queue with arrival rate $\lambda^{(\gamma)} = 2r_\gamma v\lambda$ and i.i.d. service times with distribution given in (3.21)*. Thus, $I^{(\gamma)} \sim \exp(2\lambda v r_\gamma)$ and the busy period distribution can be explicitly characterized as in [21]. Also, in the stationary regime, the probability that the SNR level crossing process is "on" is $1 - e^{-\lambda\pi r_\gamma^2}$.*

Proof. Let us consider the $M/GI/\infty$ queue defined in Theorem 1. If the queue is idle there are no nodes within a distance r_γ of the tagged user. Thus, an off period of the associated SNR level crossing process is equivalent to an idle period of the queue. Similarly, the queue's busy period is equivalent to an on period.

Since the arrival process is Poisson, the inter arrival times are exponential with parameter $\lambda^{(\gamma)} = 2\lambda v r_\gamma$. Because of the memoryless property, all idle periods have the same distribution.

Let $\hat{B}^{(\gamma)}$ denote a random variable representing the forward recurrence time associated with the on time defined as

$$\mathbb{P}(\hat{B}^{(\gamma)} > x) = \frac{1}{\mathbb{E}[B^{(\gamma)}]} \int_x^\infty \mathbb{P}(B^{(\gamma)} > z)dz. \tag{2.6}$$

Using the correspondence of the SNR level crossing process's on times with the busy period of an $M/GI/\infty$ queue, its distribution is that given in Theorem 12 in the Appendix 2.7.1 with service time $S \sim W^{(\gamma)}$ and

$$\rho^{(\gamma)} = 2r_\gamma v\lambda\mathbb{E}[W^{(\gamma)}] = \lambda\pi r_\gamma^2\ ,\ \ \nu^{(\gamma)} = 1 - e^{-\rho^{(\gamma)}}. \tag{2.7}$$

The busy period $B^{(\gamma)}$ and idle period $I^{(\gamma)}$ are independent. Hence, the SNR level crossing process is an alternating-renewal process. The mean of the busy period of the queue is given by [21]:

$$\mathbb{E}[B^{(\gamma)}] = \frac{\nu^{(\gamma)}}{2\lambda v r_\gamma (1-\nu^{(\gamma)})}. \tag{2.8}$$

Thus, the probability that the SNR level crossing process is "on" is $1 - e^{-\lambda \pi r_\gamma^2}$.

□

It is easy to see that all path loss functions which are monotonic lead to an analogue of Theorem 2.

2.3.1.2 Fluctuations of the SNR Process

The SNR process is pathwise continuous and has continuous derivatives except at a countable set of times corresponding to Voronoi cell edge crossings where the mobile sees a hand-off and where the first derivative of the SNR process is discontinuous. Since the mobile is moving at a fixed velocity v, the intensity of the cell-edge crossings is given by $4v\sqrt{\lambda}/\pi$ [22], [23].

As the mobile traverses a particular cell, there is also a time where it is the closest to the associated node. The SNR process has a local maximum at these times. Perhaps surprisingly such maxima may happen at the edge of the cell. Let us denote the point process of times corresponding to *interior* maxima by χ. The intensity of χ is given in the following theorem.

Theorem 3. *The intensity of the point process χ of the SNR maxima occurring in the interior of cells is $v\sqrt{\lambda}$.*

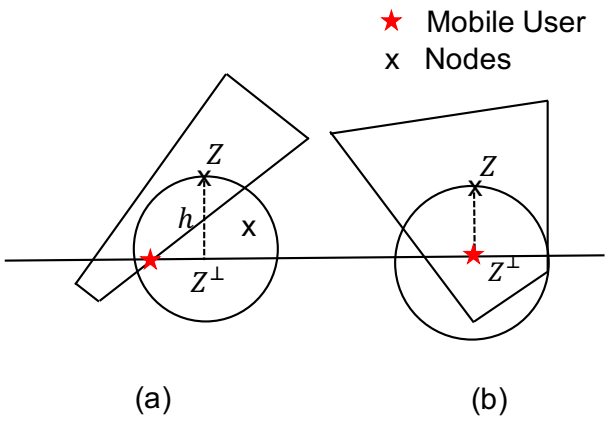


Figure 2.4: The disc $B_h(Z^{\perp})$ around the projection point $Z^{\perp}$ in two scenarios: (a) where $Z^{\perp}$ does not belong to the Voronoi cell of nucleus Z and (b) where it does.

Proof. Let $Z \in \Phi$, be a typical associated node to the mobile user at time. Let us consider the projection $Z^{\perp}$ of Z onto the straight line (mobile user's path). Then, $Z^{\perp}$, is a point of χ if $Z^{\perp}$ lies within the Voronoi cell of Z. Thus, the point process can be characterized by the fact that given a node at a height h from the straight line, the closed ball $B_h(Z^{\perp})$ of radius h centered at the projection point is empty of nodes. To that extent, consider a rectangle of unit length and height $2x$, with the mobile user's straight line path passing through the center. From Mecke's formula [24], the intensity of the point process χ is

$$\mu_{\chi} = v \lim_{x\to\infty} \lambda 2x \left(\int_{-x}^{x} e^{-\lambda\pi h^2} \frac{1}{2x} dh \right) = v \lim_{x\to\infty} \sqrt{\lambda}\mathrm{erf}(\sqrt{\lambda\pi}x) = v\sqrt{\lambda}. \quad (2.9)$$

□

Thus, the fraction of SNR maxima that happen in the interior of the

cell is given by $\pi/4$. In other words, a proportion $1 - \frac{\pi}{4}$ of the SNR maxima seen by the mobile occur at the cell edges.

2.3.1.3 Rare Events

In this subsection we consider certain rare events such as the occurrence of large SNR values. We show that the "rarity hence exponentiality" principle [7, 8] applies here. We also give the scale at which the inter-arrival times of high SNR are close to being exponentially distributed. As we shall see in Section VII, these asymptotic results can also be used for moderate values of SNR for settings typically found in wireless networks.

The following theorem gives a characterization of the asymptotic behavior of the distribution of the random variable $V^{(\gamma)}$ corresponding to the up-crossings as $\gamma \to \infty$ and as $\gamma \to 0$, which can be seen as "good" and "bad" events respectively.

Theorem 4. *For all $\gamma > 0$, the up-crossings of the SNR level crossing process, $(T_n^{(\gamma)}, n \geq 1)$, constitute a renewal process. Let $V^{(\gamma)}$ be the typical inter arrival for this process, and let $f(\gamma) = 2\lambda v r_\gamma$, then as $\gamma \to \infty$, we have*

$$f(\gamma)V^{(\gamma)} \xrightarrow{d} \exp(1),$$

and for $g(\gamma) = 2\lambda v r_\gamma e^{-\lambda\pi r_\gamma^2}$, as $\gamma \to 0$, we have

$$g(\gamma)V^{(\gamma)} \xrightarrow{d} \exp(1).$$

where $\xrightarrow{d}$ implies convergence in distribution.

Proof. The proof is given in the Appendix 2.7.2. □

The difference in the scales of the inter-event times of "good events" and "bad events" is due to the fact that the probability of occurrence of a bad event goes to zero faster than the probability of occurrence of good event. Since for bad events, as r_γ tends to ∞, a disc of radius r_γ should be empty, an event whose probability is $e^{-\lambda\pi r_\gamma^2}$. Conversely, for good events, as r_γ tends to 0, there should be at least one node in the disc of radius r_γ, an event whose probability is given by $1 - e^{-\lambda\pi r_\gamma^2}$.

2.3.2 Sharing Number Process

In contrast to the SNR or the Shannon rate process, which take their values in the continuum, the sharing number process is discrete valued and piece-wise constant. We start with evaluating the frequency of discontinuities and then introduce an upper-bound process to be used in the sequel. We conclude this section with a study of rare events.

2.3.2.1 Discontinuities

In the sequel, we will use the following simplified version of the Johnson-Mehl cell, see [22].

Definition 3. *Consider $\Phi = \{X_i\}$ the Poisson point process of nodes on $\mathbb{R}^2$. For any threshold γ, the Johnson–Mehl cell $J^{(\gamma)}(X_i)$ associated with X_i is the intersection of the Voronoi cell of X_i w.r.t. Φ with a disc of radius r_γ centered at X_i, see Fig.2.5.*

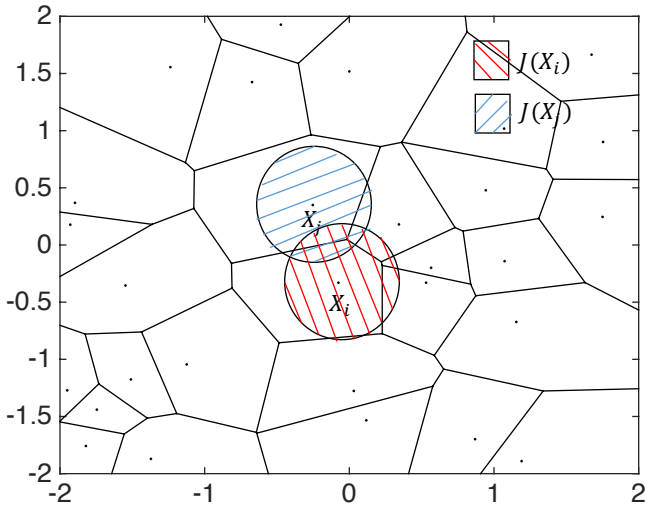


Figure 2.5: Johnson-Mehl cells.

Thus for a fixed threshold γ, the mobile user is covered/served at time t if and only if it is within the Johnson–Mehl cell of its associated node $X(t)$ for the radius r_γ. We recall the following definition.

Definition 4. *The process* $(N^{(\gamma)}(t), t \geq 0)$ *is defined as the number of static users present in the Johnson-Mehl cell* $J^{(\gamma)}(X(t))$ *if the mobile is in* $J^{(\gamma)}(X(t))$ *and 0 otherwise.*

The process $(N^{(\gamma)}(t), t \geq 0)$ is an $\mathbb{N}$-valued piece-wise constant process, with jumps at certain Johnson–Mehl cell edge crossings, see e.g., Fig 2.6. The value it assumes upon entering a cell is a conditionally Poisson random variable with a parameter that depends on the area of the cell in question. The following theorem provides an upper bound for the jump intensity:

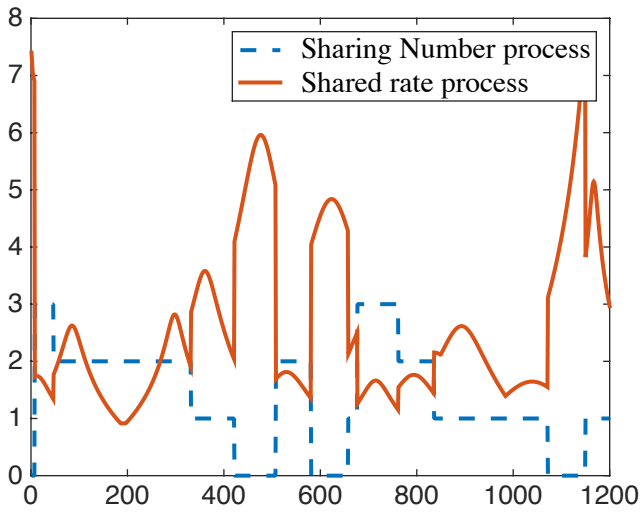


Figure 2.6: Trace of the sharing number and shared rate processes.

Theorem 5. *An upper bound for the intensity of discontinuities of the sharing number process is the intensity of the Johnson-Mehl cell edge crossings, which is given by:*

$$\frac{4v\sqrt{\lambda}}{\pi}(\text{erf}(\sqrt{\lambda\pi}r_\gamma - 2\sqrt{\lambda}r_\gamma e^{-\lambda\pi r_\gamma^2}),$$

where, $\text{erf}(z) = \frac{2}{\sqrt{\pi}}\int_0^z e^{-t^2}dt$.

Proof. The intensity of cell edge crossings in a Johnson-Mehl tessellation is given in [25]. Not all cell edge crossings lead to jumps as the number of static users can be the same across two adjacent Johnson-Mehl cells, thus the given intensity is an upper bound. □

Given that the value of $N^{(\gamma)}(t)$ is a non-zero constant, its *time of constancy* is defined as the typical amount of time it remains at that same constant. This time of constancy is lower bounded by the distribution of intersection of the motion's line with a typical Johnson–Mehl cell [25]. It is a lower bound only since the number of static users can be the same across two adjacent Johnson-Mehl cells.

2.3.2.2 Upper Bound for the Sharing Number Process

Note that the two underlying processes that are used to define the shared rate process, namely the sharing number process $(N^{(\gamma)}(t), t \geq 0)$ and the SNR process $(\text{SNR}(t), t \geq 0)$, are not independent as is easily seen in the case where $\gamma = 0$. For example, assume that the mobile user experiences a low SNR, i.e., it is far from its associated node. Then the area of the Voronoi cell is likely to be large. This in turn implies that the mobile is more likely to share its infrastructure node with a large number of users.

Fortunately, the sharing number process admits a simple upper bound which is independent of the SNR process. This upper bound is defined as follows:

Definition 5. *We define an upper bound process,*$(\hat{N}^{(\gamma)}(t), t \geq 0)$*, on the sharing number process. Let* $D(x, r_\gamma)$ *denote a disc of radius* r_γ *centered at location* x*. Let* $\hat{N}^{(\gamma)}(t), t \geq 0$*, be the random process capturing the number of static users which are in the disc* $D(X(t), r_\gamma)$ *if the mobile is in* $J^{(\gamma)}(X(t))$*, and 0 otherwise.*

The stochastic process $(\hat{N}^{(\gamma)}(t), t \geq 0)$ enjoys most of the structural properties of $(N^{(\gamma)}(t), t \geq 0)$; in particular, it is piece-wise constant and one can derive natural bounds on the intensity of its jumps. In addition and more importantly:

- For all $t \geq 0$, $\mathbb{P}(\hat{N}^{(\gamma)}(t) \geq N^{(\gamma)}(t)) = 1$, i.e., $\hat{N}^{(\gamma)}(t)$ is an upper bound for the sharing number; this bound is tight in the regime where the area of the typical Voronoi cell ($\frac{1}{\lambda}$) is large compared to the area of the disc of radius r_γ;
- When the mobile user is covered, the sharing number $\hat{N}^{(\gamma)}(t)$ is Poisson with parameter $\xi \pi r_\gamma^2$;
- The stochastic processes $(\hat{N}^{(\gamma)}(t), t \geq 0)$ and $(\mathrm{SNR}(t), t \geq 0)$ are independent.

2.3.3 Shared Shannon Rate Process

Recall that the *shared rate* is given by

$$S^{(\gamma)}(t) = R^{(\gamma)}(t) F^{(\gamma)}(t) = R^{(\gamma)}(t) \frac{1}{N^{(\gamma)}(t) + 1}, \tag{2.10}$$

where $R^{(\gamma)}(t)$ is the Shannon rate defined in (2.2) and $F^{(\gamma)}(t)$ the sharing factor at time t. A realization of this process is illustrated in Fig.2.6. In the sequel, we use our upper bound process on $N^{(\gamma)}(t)$ to obtain a lower bound on the shared rate:

$$\hat{S}^{(\gamma)}(t) = R^{(\gamma)}(t) \hat{F}^{(\gamma)}(t) = R^{(\gamma)}(t) \frac{1}{\hat{N}^{(\gamma)}(t) + 1}, \tag{2.11}$$

which is now a product of two *independent* random variables.

2.3.3.1 Shared Rate Variability

The shared rate process is continuous and differentiable except for a countable number of points and a countable set of points of discontinuity for the first derivative. It is equal to zero when the mobile is not covered. The point process of discontinuities is upper bounded by the Johnson-Mehl cell edge crossings (see Theorem 5). The point process of its local maxima where the derivative is zero is the same as that of the SNR process given in Theorem 3.

As can be seen from (2.11), the variability in the mobile user's shared rate is driven by two processes: the Shannon rate and the sharing factor process. It is of interest to understand their relative contributions. To answer this question, let us consider the variance of $\hat{S}^{(\gamma)}$ which can be written using the conditional variance formula as:

$$\mathrm{var}(\hat{S}^{(\gamma)}) = \mathrm{var}(R^{(\gamma)})\mathbb{E}[\hat{F}^{(\gamma)}]^2 + \mathrm{var}(\hat{F}^{(\gamma)})(\mathbb{E}[(R^{(\gamma)})^2]). \tag{2.12}$$

Note that if the density of users increases, the variance of the shared factor decreases whereas the variance of the Shannon rate remains constant. Thus, the variance of the shared rate varies approximately linearly with a slope equal to the second moment of the Shannon rate. By contrast if the density of nodes increases, the variance of both the sharing factor and the Shannon rate vary making their relative contributions more complex.

ξ	$\mathbf{var}(S^{(\gamma)})$	$\mathbf{var}(F^{(\gamma)})$
5	0.6527	0.0303
25	0.0674	0.0028
50	0.012	0.0016
75	0.0047	0.0006
100	0.0023	0.0001

Table 2.2: Values of variance with increasing user density, ξ, constant density of nodes, $\lambda = 25/\pi$ and constant radius of coverage $r_\gamma = 500$m.

Let us evaluate empirically what are the contributions of the user and node density to the variability of shared rate. For a given node density $\lambda = 25/\pi$ and radius of coverage $r_\gamma = 500$m, by increasing the density of static users, the variance of the shared rate decreases linearly with the variance of the sharing factor. Table 2.2 gives numerical values evaluated via simulation. For a given static user density $\xi = 50$ and radius of coverage of 200m, by increasing the density of nodes the variance of the sharing factor monotonically increases, whereas the changes in the variance of SNR process is complicated due to the way it is defined.Thus, the changes in the variance of the sharing number process is difficult to characterize with densification of nodes due to the complex relation given in (2.12). Table 2.3 gives numerical values evaluated by simulations.

2.3.3.2 Rare Events

Since the mobile user's shared rate depends on two factors it is of interest to understand their relative contributions towards the events of high/low

λ	$\mathbf{var}(S^{(\gamma)})$	$\mathbf{var}(F^{(\gamma)})$	$\mathbf{var}(R^{(\gamma)})$
5	0.0438	0.0049	0.3325
25	5.1292	0.0576	3.3944
50	0.6351	0.0642	0.9607
75	1.2626	0.0723	0.6926
100	0.9682	0.0756	0.2252

Table 2.3: Values of variance with increasing node density, λ, constant density of users, $\xi = 50$ and constant radius of coverage $r_\gamma = 200$m.

shared rate. The proofs of the following theorems are given in Appendix 2.7.3 and 2.7.4 respectively.

Theorem 6. *The likelihood of the rare events associated with high shared rate is the same (up to a logarithmic equivalence) as that for the high SNR in the sense that*

$$\lim_{s\to\infty} -\frac{1}{s}\log\left(\mathbb{P}(S^{(\gamma)} > s)\right) = \lim_{s\to\infty} -\frac{1}{s}\log\left(\mathbb{P}(\log(1+\mathrm{SNR}) > s)\right) = \frac{2}{\beta}.$$

Notice that as a direct corollary of the last theorem,

$$\begin{aligned}\lim_{s\to\infty} -\frac{1}{s}\log\left(\mathbb{P}(S^{(\gamma)} > s)\right) &= \lim_{s\to\infty} -\frac{1}{s}\log\left(\mathbb{P}(S^{(\gamma)} > s | N^{(\gamma)} = 0)\right) \\ &= \lim_{s\to\infty} -\frac{1}{s}\log\left(\mathbb{P}(S^{(\gamma)} > s, N^{(\gamma)} = 0)\right) = \frac{2}{\beta}.\end{aligned}$$

For the rare events associated with a low shared rate, we will consider the lower bound process $(\hat{S}^{(\gamma)}(t), t \geq 0)$:

Theorem 7. *The rare events associated with achieving low shared rates are predominantly the same as the rare events associated with high sharing number given that the mobile is at the covering cell edge, in the sense that for some*

sequence $\{s_n\}$ *such that* $s_n \log(1 + Kr_\gamma^{-\beta}) \in \mathbb{Z}\ \forall n$ *and* $\lim_{n\to\infty} s_n \to \infty$*, we have*

$$\lim_{n\to\infty} -\frac{1}{s_n \log(s_n)} \log\left(\mathbb{P}\left(\hat{S}^{(\gamma)} < \frac{1}{s_n}\right)\right)$$
$$= \lim_{n\to\infty} -\frac{1}{s_n \log(s_n)} \log\left(\mathbb{P}\left(\frac{\log(1 + Kr_\gamma^{-\beta})}{\hat{N}^{(\gamma)} + 1} < \frac{1}{s_n}\right)\right) = \log(1 + Kr_\gamma^{-\beta}).$$

Theorem 8. *Conditioned on a very good or very bad shared rate, the relative contribution of the mobile user's location and the network congestion is as follows:*

$$\lim_{s\to\infty} \mathbb{P}(R^{(\gamma)} > s, N^{(\gamma)} = 0 \mid S^{(\gamma)} > s) = 1, \tag{2.13}$$

$$\mathbb{P}(N^{(\gamma)} > as \log(1 + \gamma) - 1 \mid 0 < S^{(\gamma)} < 1/s) = 1. \tag{2.14}$$

Conditioned on a very high shared rate, the mobile user has to be close to the associated node with no other static users sharing its resources. By contrast, conditioned on the mobile being served and experiencing a very low shared rate, it has to share its associated node with a number of users that is inversely proportional to the desired shared rate.

From the previous theorem, we know that high shared rates are predominantly the same as the event where the SNR is high and the sharing number is zero. Thus, for a given threshold δ for the shared rate process, we can study the asymptotic behavior of the distribution of the inter arrival time of shared rate up-crossings as $\delta \to \infty$.

Corollary 1. *Let* $\hat{Z}_\delta$ *denote the inter arrival time for the good events of the lower bound process* $(\hat{S}^{(\gamma)}(t), t \geq 0)$ *and let* $f(r_\delta) = 2\lambda v r_\delta$*, then as* $\delta \to 0$*, we*

have

$$e^{-\xi\pi r_\gamma^2}\hat{Z}_\delta \xrightarrow{d} \exp(1).$$

Proof. The result follows from the fact that $\hat{Z}_\delta = \sum_{i=1}^{X} V^{r_\delta}$, where X is a geometric random variable with parameter $e^{-\xi\pi r_\gamma^2}$ and $V^{(r_\delta)}$ the typical interval time of the renewal process of up-crossings associated with the SNR process of threshold δ. □

2.4 Simulations and Validation

In this section we evaluate the robustness of our mathematical model and associated asymptotic results to more realistic settings. We use simulation to study the temporal variations of the SNR process experienced by a mobile user under various scenarios which are not captured by our analytical framework. Our model is challenged in various complementary ways: e.g., by adding fading, accounting for interference from other base stations and relaxing the straight line mobility model. In each case the objective is to determine to what degree our simplified mathematical model is still approximately valid, providing robust engineering rules of thumb to predict what mobile users see in practice. In particular, we will answer the following questions:

- How quickly do the SNR up crossings converge to exponential asymptotics as a function of the associated thresholds?
- Are the results obtained robust to the presence of fast fading?

- Are there regimes where the temporal characteristics of the SNR process are good proxies for those of the SINR process, e.g., high path loss?
- Does the characterization of the SNR level crossing process valid for other mobility models?

We begin by introducing our simulation methodology and the default parameters used throughout this section.

2.4.1 Simulation Methodology

We initially consider a user moving on a straight line (road) at a fixed velocity of 16 m/s. The base stations are randomly placed according to a Poisson point process with intensity λ such that the mean coverage area per base station is equal to that of a disc with radius 200m. Unless otherwise specified, we consider the path loss function given by Eq. (2.1) with exponent $\beta = 4$ and assume that all base stations transmit with equal power $p = 2$W. The signal strength received by the mobile user is recomputed every 10^{-2} seconds. The total distance traveled by the mobile is 750 km along a straight line. The number of simulation runs considered to evaluate the means is 1000. The number of handovers experienced by the mobile while traveling this distance is on average 4269.

We calibrate the thermal noise power to the *cell-edge* user. Let D be a random variable denoting the distance of a typical user to the closest base station. Define d_{edge} by the relation $P(D \leq d_{edge}) = 0.9$. Since in our

simulation setting $P(D \leq d) = 1 - \exp(\lambda \pi d^2)$, we have $d_{edge} = \sqrt{\frac{-\ln(0.1)}{\pi \lambda}}$. If we fix the desired SNR at the cell edge to be SNR_{edge}, this then determines the noise power to be $w = \frac{p d_{edge}^{-\beta}}{\text{SNR}_{edge}}$. We set the SNR_{edge} to be 1 dB.

In the sequel we evaluate how quickly the convergence to exponential stated in Theorem 4 arises. To that end we compare the re-normalized distributions obtained via simulation to a reference exponential distribution with parameter 1 using the Kolmogorov-Smirnov (K-S) test. The K-S test finds the greatest discrepancy between the observed and expected cumulative frequencies– called the "D-statistic". This is compared against the critical D-statistic for that sample size with 5% significance level. If the calculated D-statistic is less than the critical one, we conclude that the distribution is of the expected form, see e.g. [26].

2.4.2 Convergence of Level-crossing Asymptotics

Theorem 4 shows that as the SNR threshold γ in dB increases, the rescaled distribution for up-crossings of the SNR process becomes exponential. The question is how large γ needs to be for this result to hold. To that end we simulated the level crossing process for various γ and subjected it to the K-S test as mentioned above.

The Q-Q plots for up-crossing inter-arrivals rescaled by $f(r_\gamma)$ as introduced in the theorem are depicted in Fig. 2.7. As expected, as the threshold increases, the points become linear, and for a threshold value of $\gamma = 50$ or more, it is accepted as an exponential with unit mean by the K-S test.

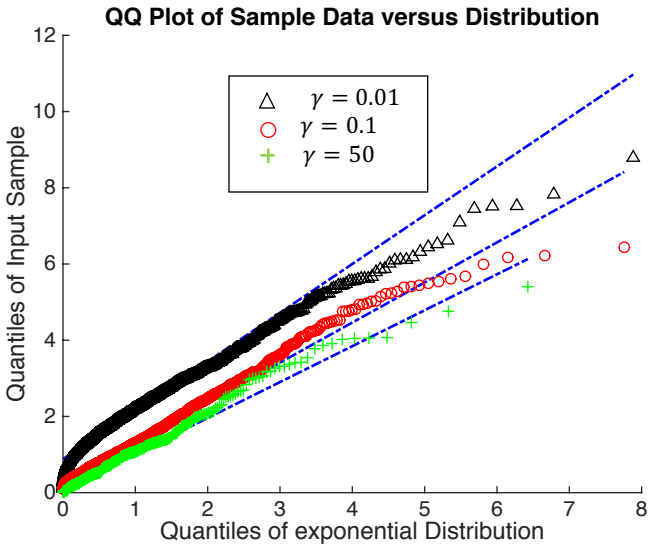


Figure 2.7: Q-Q plot for various thresholds.

In practice a SNR of 50 dB is not realistic for wireless users. However, as seen from the Q-Q plots in Fig.2.7, for moderate values of γ such as 0.1, the up-crossing inter-arrivals can be well approximated by an exponential with parameter $1/f(r_\gamma)$. Further, the absolute error in the mean of the inter-arrival of up-crossings for moderate value of $\gamma = 1$ is approximately only 9.3 s.

2.4.3 Robustness of Level-crossing Asymptotics to Fading

Next we study the effect that channel fading might have on the level-crossing asymptotics. We consider channels with Rayleigh fading with unit mean, so that the SNR experienced by the tagged mobile user at a distance d from the base station is given by $Hpd^{-\beta}/w$, where H is a fading random variable which is exponential with unit mean. The coherence time is set to

$t_c = 0.423/f_d$, where f_d is the Doppler shift given by $f_d = \frac{v}{c} f_o$ where v is the vehicle velocity, c is the speed of light and $f_o = 900$MHz is the operating frequency. This gives a coherence time $t_c = 0.007s$. Thus, fading (power) changes every 0.007 seconds. A realization of the SNR process with fading is exhibited in Fig. 2.8.

If we assume no fast fading, the realizations of the SNR process seen by a mobile are continuous and differentiable and level crossings are well defined. However when the fluctuations associated with fast fading are added to the model, level crossings will exhibit bursts of up-crossings associated with fast fading when the mean channel gains reach the thresholds, see Fig. 2.8. Thus, in our simulations, we avoid these bursts of up-crossings by considering the first up-crossing and by suppressing all subsequent up crossings (if any) for an appropriate time period. We take a time period for the suppression of up-crossings equal to twice the expected on time of $2E[B^{(r_\gamma)}]$ [21].

In order to vary the variance while keeping the mean of the fading equal to 1, we now consider fading which is a mixture of exponentials. For this process, we would expect that for fading with mean one, if variance is small, the appropriately rescaled inter-arrival distribution for up-crossings which are not suppressed would once again be asymptotically exponential with parameter 1. In other words we expect geometric variations associated with base station locations to dominate channel variations. Whereas, if the fading variance is high, one might expect the SNR threshold required to see convergence to an exponential to increase. Fig. 2.9 exhibits such thresholds as a function of the

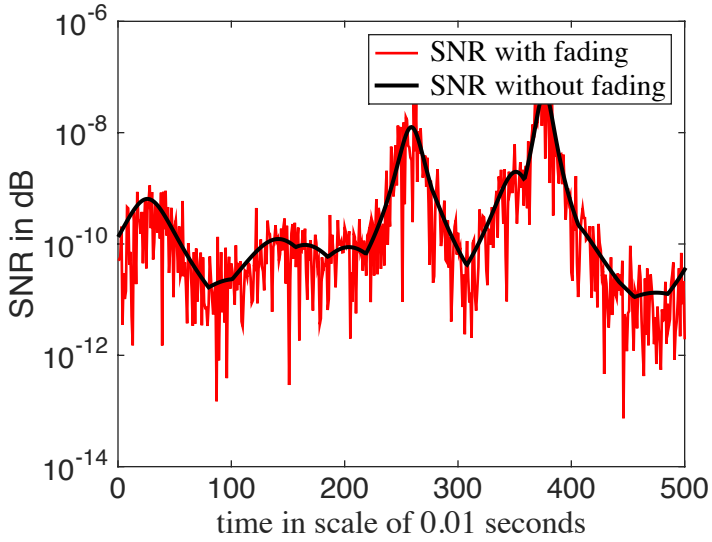


Figure 2.8: SNR process in presence of fading with mean 1.

fading variance. As can be seen for fading variances exceeding 8, the channel variations dominate the geometric variations and thus up-crossing asymptotics differ from those of Theorem 4.

2.4.4 Robustness of Level-crossing Asymptotics to Interference

So far we have focused on the SNR process. One might ask to what degree the Signal-to-Interference-plus-Noise Ratio (SINR) process, share similar characteristics.

To that end we simulated the SINR process for a setting with a high path loss exponent of $\beta = 4$ and found once again that the rescaled distribution for the up-crossing inter-arrivals converges to an exponential with parameter 1. The test requires a threshold $\gamma = 31.7$dB.We then evaluated, for different path

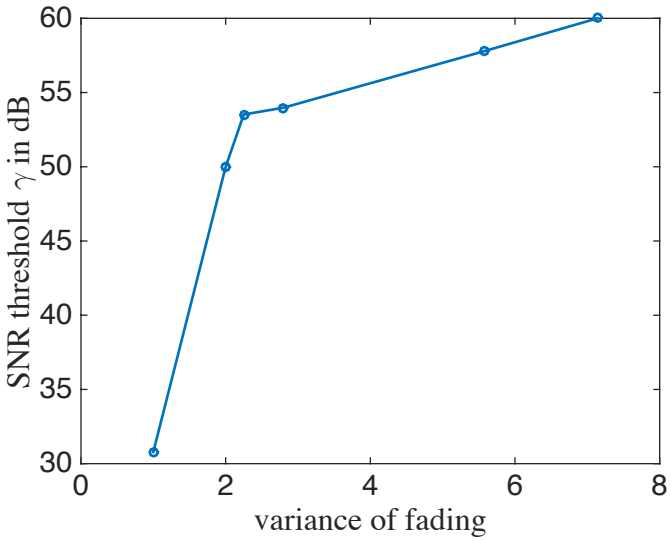


Figure 2.9: Threshold above which the inter-arrival of up-crossing converges to an exponential with parameter 1, for different variance of fading.

loss β, what threshold values are needed to obtain a similar convergence. As shown in Fig. 2.10, the threshold in question increases as β decreases. Further we found that for $\beta < 3.5$, we no longer have the desired convergence property. In summary, for high path-loss exponents $\beta \in [3.5, 4]$, the up-crossing asymptotics for the SNR and SINR processes are similar.

2.4.4.1 Robustness to Generalized Mobility Models

far we have focused on a mobile moving strictly along a straight line. Next we evaluate to what extent one can relax this requirement. To that end we consider a variant of the random way point model, [27], where a user moves at a constant velocity, v meters per sec, but every t_0 secs, with probability p_0 selects a new independent direction (angle) uniformly distributed in $(-\theta_0, \theta_0)$.

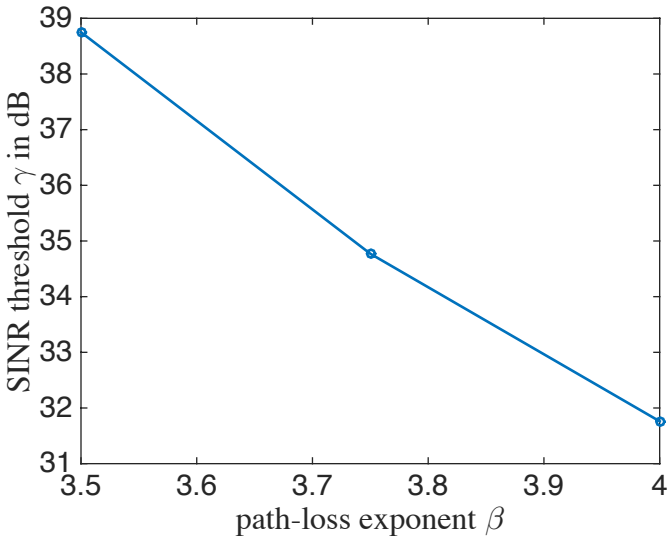


Figure 2.10: Threshold above which the inter-arrival of up-crossing converges to an exponential with parameter 1, for different path loss exponent.

We conducted simulations where for a given SNR threshold,γ, various values for the parameters, p_0, θ_0, v and t_0 were varied and the following quantities were measured:

- The means of the on-off periods;
- The independence of the on-off periods;
- The distribution of the up-crossing inter-arrivals.

We considered three particular sets of values for the parameters (v, t_0, p_0, θ_0), each reflecting a specific kind of traffic: (1) Car on a highway: $(20, 10, 0.001, \pi/45)$, (2) Car in an urban area: $(10, 1, 0.1, \pi/4)$ and (3) Wanderer: a pedestrian moving in a city: $(1, 0.1, 0.6, \pi)$. Fig.2.11 illustrates a snapshot of the paths of the

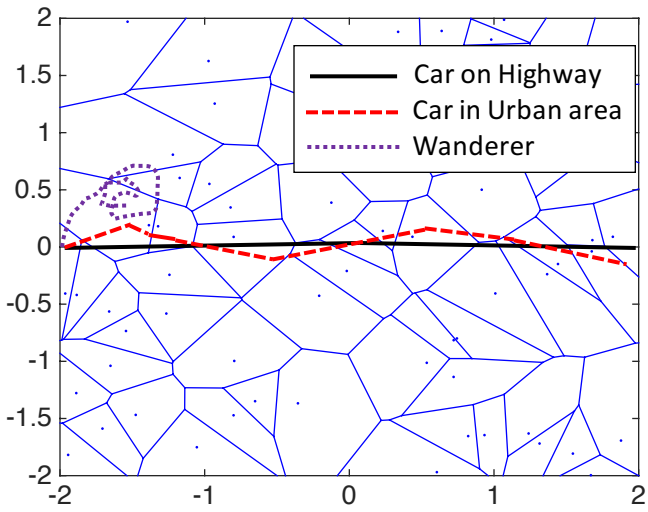


Figure 2.11: Three specific mobile user motion paths.

three cases.

The mean of the on-off periods of the SNR process is within a 5 percent error margin in the case (1) but they do not match for the other two cases for moderate values of the SNR threshold. However, for certain high threshold values, $\gamma \in [20\text{dB}, 60\text{dB}]$ the mean coverage time for the case (2), is within an acceptable error margin of 10 percent. Further, the mean on time for both the second and third cases is larger than the mean on time in the case where the mobile is moving along a straight line.

The on-off periods are also highly correlated for the later two cases, while independence approximately holds in case (1). The distribution of the up-crossing inter-arrivals for a specific high threshold value of $\gamma = 50$ dB is now considered. The convergence to an exponential of unit mean is only

observed for the first case. In summary, the robustness of our model is only guaranteed for slight changes in the direction. Major changes in the direction have a strong effect on the characterization of on-off periods as an alternating renewal process with the proposed distribution.

2.5 Applications

In this section we use our framework to evaluate application-level performance of mobiles in a shared wireless network. In particular we consider two very different wireless scenarios: (1) video streaming to mobiles sharing a cellular or WiFi network and (2) large file downloads using WiFi. In both applications, we consider a service model where the users are served only if the SNR experienced by them is above a certain threshold. Thus, the users are served only during the on period of the SNR process. We illustrate how the characterization of the level-crossing process can be leveraged to improve the QoE of the tagged user.

2.5.0.1 Stored Video Streaming to Mobile Users

Let us consider a scenario where mobile users are viewing stored videos which are streamed over a sequence of wireless downlinks. The users are distributed according to a Poisson point process of density ξ and are moving independently of each other. Consider a policy where a mobile user is served by a node only if the SNR experienced is greater than a threshold γ. Thus, the nodes serve the mobile users within the radius of coverage $r_\gamma = (p/(\gamma w))^{1/\beta}$.

A lower threshold γ corresponds to a lower transmission rate (when served), a higher probability of coverage and sharing with a large number of other mobile users. Conversely a higher threshold implies higher transmission rate and sharing with fewer other mobile users. For simplicity we consider rebuffering as the primary metric for user's video QoE [20].

The playback buffer state of the tagged mobile user moving at a fixed velocity v on a straight line, can be modeled as a fluid queue.The arrival rate to the queue alternates between an average ergodic rate, $h(\gamma) = \mathbb{E}[R^{(\gamma)}|\text{SNR} > \gamma]$ and zero depending on whether the mobile is being served or not. Let η denote the video playback rate in bits. Hence, as long as the buffer is non-empty, the fluid depletion rate of the queue is η. Re-buffering of the video is directly linked to the proportion of time the playback buffer is empty, which is given by $1 - \rho(\gamma, \xi)$, where the *load factor*, $\rho(\gamma, \xi)$, [24] of the queue is:

$$\rho(\gamma, \xi) = \frac{\nu^{(\gamma)} h(\gamma)}{\eta} \mathbb{E}\left[\frac{1}{\hat{N}_p^{(\gamma)} + 1} | \text{SNR} > \gamma\right]. \tag{2.15}$$

where $\nu^{(\gamma)}$ is the probability that the alternating renewal process associated with the arrival rate to the fluid queue is "on".

The first natural question one can ask is whether there is a choice of γ such that the fluid queue is unstable, thus ensuring no rebuffering in the long term. In other words, does there exist a $\gamma > 0$ such that $\rho(\gamma, \xi) > 1$? Given our policy and the QoE metric, the network provider may choose a lower value of threshold,γ, as long as the typical mobile user in the long run experiences

no rebuffering.

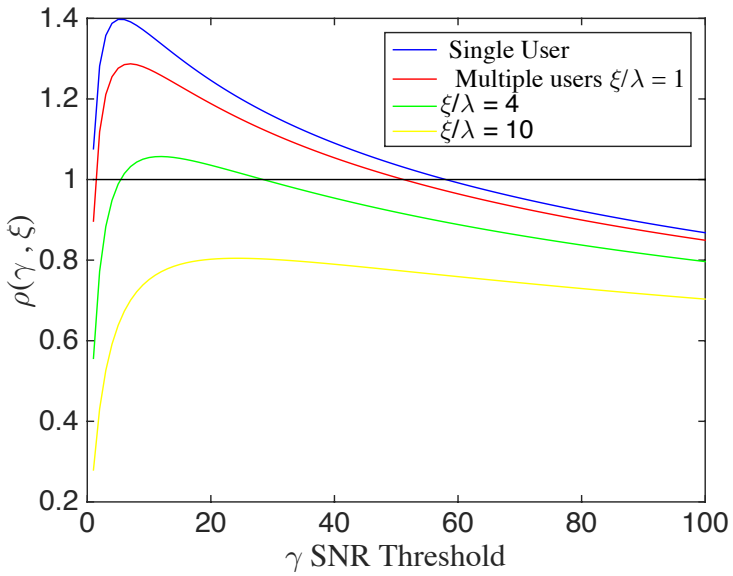


Figure 2.12: The load factor of the fluid queue for a single user (top curve) and for a positive density of users as a function of γ (curves below the top curve). For the latter curves, the number of users per base station are $\xi/\lambda = 1, 4$ and 10 from top to bottom. Here $b = 1$. All functions can be multiplied by an arbitrary positive constant when playing with a and η.

Now, for simplicity let us consider a constant mean transmission rate $\kappa = a \log(1+\gamma)\, \mathbb{E}\left[\frac{1}{N_p^{(\gamma)}+1}\right]$ instead of the average ergodic rate. This constant bit rate is when the network does not rely on adaptive coding/decoding. For additional motivation for this scenario, see [20]. The load factor $\rho(\gamma, \xi)$ is then given by

$$\rho(\gamma, \xi) = \frac{a \log(1+\gamma)\left(1 - e^{\frac{-b}{\gamma^{2/\beta}}}\right)}{\eta} \mathbb{E}\left[\frac{1}{N_p^{(\gamma)}+1}\right], \tag{2.16}$$

where $b = \lambda\pi \left(\frac{p}{w}\right)^{\frac{2}{\beta}}$ and $\mathbb{E}[1/(N_p^{(\gamma)} + 1)]$ can be calculated by numerical integration as described in previous section.

It is easy to check that the function $\rho(\gamma, \xi)$ has a unique maximum γ^* on $(0, \infty)$. A plot of (2.16) and the value of γ^* are exhibited in Fig. 2.12. Notice that for these parameters, the value of γ^* increases with ξ.

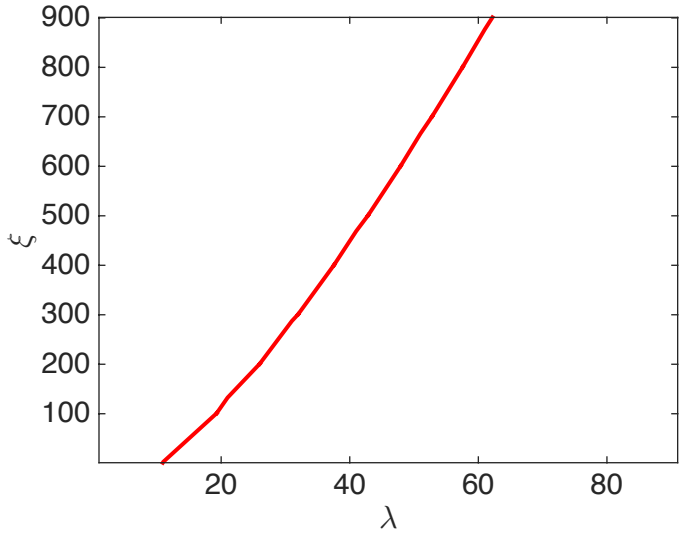


Figure 2.13: Level set curve of $\rho(\gamma^*, \xi) = 1$ for an arbitrary positive constants a and η.

For a given base station density λ and density of users ξ, one can evaluate the SNR threshold value γ^* for which the load factor ρ is maximum. Fig.2.13 illustrates the level set curve of $\rho(\gamma^*, \xi) = 1$ for various values of λ and ξ. In this setup, given the video consumption rate η, it is possible to answer questions like what is the minimum density of base stations required to serve a certain density of users such that in the long term the video streaming is

uninterrupted for all the users.

Remark 1. *In the case where there exists no threshold γ for which the condition for no long term rebuffering is satisfied and $\kappa > \eta$, the fluid queue alternates between busy and idle period representing the periods when the video is un-interrupted or frozen, respectively. The distribution of the on periods $B^{(\gamma)}$ and that of the off periods $I^{(\gamma)}$ of the $M/GI/\infty$ queue discussed in Section determine the distribution of the busy period B_f and that of the idle period I_f of the fluid queue. When denoting by κ the constant input rate during on period and by η the constant output rate when the queue is non-empty, the Laplace transform of B_f is given by [28]*

$$\mathcal{L}_{B_f}(s) = \mathcal{L}_{B^{(\gamma)}}(s\sigma + \lambda(\sigma - 1)(1 - \mathcal{L}_{B_f}(s))), \tag{2.17}$$

where $\sigma = \kappa/\eta > 1$, and the idle period I_f has an exp$(\lambda^{(\gamma)})$ distribution.

2.5.0.2 WiFi Offloading

WiFi offloading helps to improve spectrum efficiency and reduce cellular network congestion. One version of this scheme is to have mobile users opportunistically obtain data through WiFi rather than through the cellular network. Offloading traffic through WiFi has been shown to be an effective way to reduce the traffic on the cellular network, when available WiFi is faster and uses less energy to transmit the data.

Let us consider a scenario where the mobile users download a large file from the service provider, relying on WiFi hotspots, distributed according

to a Poisson point process of intensity λ, rather than from the cellular base stations. The users are distributed according to a Poisson point process of density ξ and are moving independently of each other. Assume that the WiFi hotspots have a fixed coverage area, i.e., the mobile user connects to WiFi only if it is within a certain distance r from the hotspot. Thus, higher the density of hotspots, λ, deployed by the provider the better the performance experienced by mobile users which rely only on them. We consider the time it takes to complete the download as the primary metric for user's quality of experience.

Consider again the case without adaptive coding/decoding and the tagged mobile user moving on a straight line at constant velocity v. Then, the shared rate experienced by the mobile user is the constant

$$\kappa = a \log(1+\gamma)\, \mathbb{E}\left[\frac{1}{N_p^{(\gamma)}+1}\right]$$

as defined above. In addition, the mobile user experiences an alternating on and off process, as characterized in Theorem 2.

Below, for the sake of mathematical simplicity, we assume that the file size F is exponential with parameter δ and that the mobile user starts to download the file at the beginning of an on period. Let T be a random variable denoting the time taken to download the file. Consider the event $J = \{F > \kappa B^{(\lambda)}\}$, where $B^{(\lambda)}$ is a random variable distributed like a typical on period seen by the mobile (see Theorem 2) independent of F and let

$$\alpha = \mathbb{P}(J) = \mathbb{P}(F > \kappa B^{(\lambda)}) = \mathcal{L}_{B^{(\lambda)}}(\delta\kappa). \tag{2.18}$$

Now, define the non-negative random variable X as the busy period, $B^{(\lambda)}$, conditioned on event J, whose c.d.f. is given by:

$$\mathbb{P}(X < x) = \frac{1}{\alpha} \int_0^x e^{-\delta\kappa z} f_{B^{(\lambda)}}(z) dz \tag{2.19}$$

Conditioned on the event, $\{F \leq \kappa B^{(\lambda)}\}$, let Y denote the time taken to download the file, i.e., $\frac{F}{\kappa}$. The c.d.f. of the non-negative random variable Y is then given by:

$$\mathbb{P}(Y < y) = \frac{1}{1-\alpha} \int_0^{y\kappa} \delta e^{-\delta z} \mathbb{P}(B^{(\lambda)} > \frac{z}{\kappa}) dz. \tag{2.20}$$

Notice that

$$\mathcal{L}_X(s) = \frac{1}{\alpha} \mathcal{L}_{B^{(\lambda)}}(s + \delta\kappa) \tag{2.21}$$

$$\mathcal{L}_Y(s) = \frac{1}{1-\alpha} \int_0^{\infty} \delta\kappa e^{-y(\delta\kappa+s)} \mathbb{P}(B^{(\lambda)} > y) dy. \tag{2.22}$$

The following representation of the Laplace transform of T is an immediate corollary of the on-off structure:

Theorem 9. *Consider a network of WiFi hotspots distributed according to a Poisson point process of intensity λ and radius of coverage r shared by mobile users distributed independently according to a Poisson point process of intensity ξ. Assuming that the tagged mobile user starts to download the file at the beginning of an on period, the Laplace transform of the time taken to download a file of size $F \sim \exp(\delta)$ is given by*

$$\mathcal{L}_T(s) = \frac{(1-\alpha)\mathcal{L}_Y(s)}{1 - \alpha\mathcal{L}_X(s)\frac{2\lambda v r}{2\lambda v r + s}}. \tag{2.23}$$

Proof. Proof is given in the Appendix 2.7.7. □

It is remarkable that the Laplace transform of T admits a quite simple expression in terms of that of $B^{(\lambda)}$. Other and more general file distributions can be handled as well when using classical tools of Laplace transform theory. Note that this setting also leads to interesting optimization questions such as the optimal density of WiFi hotspots needed to be deployed for lower expected download times.

2.6 Variants

In this section, we consider two generalizations of our framework: (1) we move from sharing with static users to sharing with mobile users, and (2) we move from homogeneous to heterogeneous wireless infrastructures.

2.6.1 Sharing the Network with Mobile Users

Until here we have considered a tagged mobile user sharing the network with static users. We now consider two scenarios: (1) that where the other users are mobile and are initially distributed according to a homogeneous Poisson point process, and (2) that where they are mobile but restricted to a random road network, i.e., form a Cox process.

2.6.1.1 Homogeneous Poisson case

Consider the case where other users sharing the network are initially located according to a homogeneous Poisson point process of intensity ξ, and

subsequently exhibit arbitrary independent motion. This is a possible model for pedestrian motion. It follows from the displacement theorem for Poisson point processes [14] that the other users at any time instant will remain a Poisson point process of intensity ξ.

As considered before, the users are served only if they are at a distance less than r_γ from the closest node. Consider a tagged mobile user moving at a fixed velocity along a straight line. Let us define the mobile sharing number process $(N_p^{(\gamma)}(t), t \geq 0)$ as the number of users sharing the node associated with the tagged user when it is served and zero otherwise. Thus, at any given time t, the tagged user shares its resources with a random number of users $N_p^{(\gamma)}(t)$ which is Poisson with a parameter depending on the area of the Johnson-Mehl cell.

For simplicity, we define the shared rate process $(S_c^{(\gamma)}(t), t \geq 0)$ as

$$S_c^{(\gamma)}(t) = \begin{cases} a \log(1+\gamma)\, \mathbb{E}\left[\frac{1}{N_p^{(\gamma)}+1}\right] & \text{if } \; L(t) \leq r_\gamma, \\ 0 & \text{otherwise.} \end{cases} \tag{2.24}$$

Since the distribution of the area of the Voronoi cell is unknown, we approximate $N_p^{(\gamma)}$ to be Poisson with parameter $\xi \mathbb{E}[\hat{J}^{(\gamma)}]$, where $\hat{J}^{(\gamma)}$ denotes the area of the Johnson-Mehl cell of radius r_γ, conditioned on the fact that the tagged user is within a distance r_γ from the associated node, which introduces an additional bias. We can compute the expectation with the help of integral geometry as shown in Appendix 2.7.5.

We found the value of the expectation using numerical integration and

compared the mean number of users $\mathbb{E}[N_p^{(\gamma)}] = \xi\mathbb{E}[\hat{J}^{(\gamma)}]$ with the sample mean obtained from simulation for various parameters λ, r_γ. To validate the approximation of $N_p^{(\gamma)}$ by a Poisson random variable with parameter $\xi\mathbb{E}[\hat{J}^{(\gamma)}]$, we compared the value of $\mathbb{E}[1/(N_p^{(\gamma)}+1)]$ calculated using numerical integration with that from simulations. We found that the calculated value of the expectation is within the 95% confidence interval of the simulated mean.

2.6.1.2 Cox process

Let us consider a population model where roads are distributed according to a Poisson line process of intensity λ_r on $\mathbb{R}^2$ [29]. Then independently on each road, we consider users distributed according to a stationary Poisson point process of intensity λ_t. This is known as a Cox process and we denote it by Φ_u [29]. This model can be used to represent a car motion on a road network.

For a line L of the line process, let us denote the orthogonal projection of the origin O on L by (θ, r) in polar coordinates. For $\theta \in [0, \pi)$ and $r \in \mathbb{R}$, (θ, r) is unique. Thus, a Poisson line process with intensity λ_r is the image of a Poisson point process with the same intensity on half-cylinder $[0, \pi) \times \mathbb{R}$.

Suppose all users on the roads are moving arbitrarily but independently from each other. Thus, at any instant the distribution of users on a given road remains Poisson. Consider a tagged user moving along a given road, then we have the following theorem by [30].

Theorem 10. *Φ_u is stationary, isotropic, with intensity $\pi\lambda_r\lambda_t$. From the*

point of view of the tagged user i.e., under its Palm distribution, the point process is the union of three counting measures: (1) the atom at the origin, O, (2) an independent λ_t-Poisson point process on a line through O with a uniform independent angle and (3) the stationary counting measure Φ_u.

Following our previous framework, let us define an another sharing number process $(N_d^{(\gamma)}(t), t \geq 0)$ as the number of users sharing the node associated with the tagged user when it is served and zero otherwise.

Evaluation of the mean of $N_d^{(\gamma)}$. Suppose the tagged user is at the origin O. Let the associated node $X(t)$ be at a distance x from the origin. Let $d(0, \theta)$ be a line through the origin with θ uniform from $[0, \pi)$. From the aforementioned theorem, the number of sharing users $N_d^{(\gamma)}$ can be split into two terms: N_s denoting the number of sharing users from stationary Φ_u and N_l denoting the number of sharing users on the line $d(0, \theta)$.

Given any convex body Z, $N_s(Z)$ denotes the number of sharing users from stationary Φ_u present in Z, and $\mathbb{E}[N_s(Z)]$ is given by $\pi\lambda_r\lambda_t \text{area}(Z)$ [31]. Let $\hat{J}^{(\gamma)}$ denote the area as defined before. Thus,

$$\mathbb{E}[N_s] = \pi\lambda_r\lambda_t\mathbb{E}[\hat{J}^{(\gamma)}],$$

where, $\mathbb{E}[\hat{J}^{(\gamma)}]$ is evaluated using integral geometry (see Appendix 2.7.5).

Now, let $l(0, \theta)$ denote the length of the line $d(0, \theta)$ in the area $\hat{J}^{(\gamma)}$. Then, N_l, the number of sharing users on the line $d(0, \theta)$, is Poisson with parameter $\lambda_t l(0, \theta)$. Thus, $\mathbb{E}[N_l] = \lambda_t\mathbb{E}[l(0, \theta)]$.

One can evaluate $\mathbb{E}[l(0,\theta)]$ using integral geometry (see Appendix 2.7.6). We have,

$$\mathbb{E}[N_d^{(\gamma)}] = \mathbb{E}[N_s] + \mathbb{E}[N_l] = \pi\lambda_r\lambda_t\mathbb{E}[\hat{J}^{(\gamma)}] + \lambda_t\mathbb{E}[l(0,\theta)]. \tag{2.25}$$

Thus, the mean number of users sharing the tagged user's association node is larger when the users are distributed according to a Cox process than when the users are distributed according to a Poisson point process, assuming that both have the same mean spatial intensity.

2.6.2 Mixture of Pedestrian and Road Network

Suppose now we have two types of users : drivers who stay on roads, and pedestrians which are unconstrained. If pedestrians are supposed to follow a Poisson point process, from their point of view, the number of sharing users corresponds to the sum of a stationary Poisson point process and a stationary Poisson line process. On the other hand, from a driver's point of view, the number of sharing users corresponds to the same sum, but in addition with a Poisson point process on a road the driver is taking. Thus, the mean number of sharing users is always greater for drivers. In other words, pedestrians are likely to share their infrastructure node with fewer users than drivers.

2.6.3 Heterogeneous Networks

Let us consider a deployment of micro-base stations $\hat{\Phi} = \{\hat{X}_1, \hat{X}_2, ..\}$ distributed according to some homogeneous Poisson process of intensity $\hat{\lambda}$ independent of the existing macro-base stations $\Phi = \{X_1, X_2, ..\}$. Assume

that all micro-BS transmit at a fixed power $\hat{p}$.

Let us consider a mobile user moving at a fixed velocity v along a straight line. For a given SNR threshold γ, the mobile is served by a micro-base station if its distance from its closest micro-BS is less than $\hat{r}_\gamma = (\hat{p}/w\gamma)^{1/\beta}$. Otherwise it is served by a macro-base station provided its distance from the closest macro-BS is less than r_γ.

Note that the SNR level crossing process as previously defined is again an alternating renewal process which now depends on the heterogeneous resource deployment.

Theorem 11. *For heterogeneous networks with preferential association to micro base stations, the probability that the stationary SNR level crossing process seen by a tagged user is "on" is* $1-e^{-\pi(\lambda r_\gamma^2+\hat{\lambda}\hat{r}_\gamma^2)}$. *Also, the mean time for which the process is "on" is given by*

$$\frac{e^{\pi(\lambda r_\gamma^2+\hat{\lambda}\hat{r}_\gamma^2)}-1}{2v(\lambda r_\gamma+\hat{\lambda}\hat{r}_\gamma)}. \tag{2.26}$$

Proof. In order to characterize the SNR level crossing process, we establish a connection to a Boolean model. Assume that the mobile user is moving with unit velocity. Let $B(X_i, r_\gamma)$ denote the closed ball of radius r_γ centered at X_i and $B(\hat{X}_i, \hat{r}_\gamma)$ a closed ball of radius $\hat{r}_\gamma$ centered at $\hat{X}_i$. The union of all these closed balls forms a Boolean model

$$\mathcal{E} = \left(\cup_{X_i \in \Phi} B(X_i, r_\gamma) \right) \cup \left(\cup_{\hat{X}_i \in \hat{\Phi}} B(\hat{X}_i, \hat{r}_\gamma) \right). \tag{2.27}$$

Now, assume that $\mathcal{E}$ is intersected by the directed line $\vec{l}$. Note that the Boolean model under consideration has independent convex grains and thus the intersection $\mathcal{E} \cap \vec{l}$ yields an alternating sequence of "on" and "off" periods which are independent. Let $B_h^{(\gamma)}$ and $I_h^{(\gamma)}$ be random variables denoting the length of a typical on and off periods respectively.

The distribution of the length of off period $I_h^{(\gamma)}$ is easy to establish using the contact distribution functions and is exponential with parameter λ^* [22]:

$$f_{I_h^{(\gamma)}}(l) = \lambda^* e^{-\lambda^* l},$$

where, λ^* is $2(\lambda + \hat{\lambda})\mathbb{E}[R_h^{(\gamma)}]$. Here $R_h^{(\gamma)}$ is the random variable denoting the radius of the closed ball and is given by

$$R_h^{(\gamma)} = \begin{cases} r_\gamma \text{ w.p. } \frac{\lambda}{\lambda+\hat{\lambda}}, \\ \hat{r}_\gamma \text{ w.p. } \frac{\hat{\lambda}}{\lambda+\hat{\lambda}}. \end{cases} \tag{2.28}$$

Thus, the mean off period under the assumption of unit velocity is

$$\mathbb{E}[I_h^{(\gamma)}] = \frac{1}{\lambda^*} = \frac{1}{2(\lambda r_\gamma + \hat{\lambda}\hat{r}_\gamma)}.$$

Now, in the stationary regime, the probability that the mobile user is either within a distance $\hat{r}_\gamma$ from a micro-BS or a distance r_γ from a macro-BS i.e., it's SNR level crossing process is "on", is given by the volume fraction [22]:

$$c = 1 - e^{(-(\lambda+\hat{\lambda})\bar{V})}, \tag{2.29}$$

where

$$\bar{V} = \pi\mathbb{E}[(R_h^{(\gamma)})^2] = \frac{\pi(\lambda r_\gamma^2 + \hat{\lambda}\hat{r}_\gamma^2)}{\lambda + \hat{\lambda}}. \tag{2.30}$$

Thus, $c = 1 - e^{-\pi(\lambda r_\gamma^2 + \hat{\lambda}\hat{r}_\gamma^2)}$. The probability evaluated above does not depend on the velocity of the mobile user and thus holds for any constant velocity v. The mean on period $\mathbb{E}[B_h^{(\gamma)}]$ can be evaluated using the following relation

$$c = \frac{\mathbb{E}[B_h^{(\gamma)}]}{\mathbb{E}[B_h^{(\gamma)}] + \mathbb{E}[I_h^{(\gamma)}]},$$

which results in

$$\mathbb{E}[B_h^{(\gamma)}] = \frac{e^{\pi(\lambda r_\gamma^2 + \hat{\lambda}\hat{r}_\gamma^2)} - 1}{2(\lambda r_\gamma + \hat{\lambda}\hat{r}_\gamma)}.$$

Since, the mobile user is moving with a fixed velocity v, the mean time for which the mobile user is "on" is given by $\frac{\mathbb{E}[B_h^{(\gamma)}]}{v}$.

□

These results provide an analytical characterization of the impact of heterogeneous densification on the mobile user's temporal performance. Let us now compare the performance improvement seen by the mobile user in a heterogeneous network as compared to that of a homogeneous network. The graphs in Fig. 2.14 and Fig. 2.15 illustrate the difference in the expected on-times and volume fraction of the networks respectively.

Notice that as micro-base stations are added, the expected on time decreases and later increases. The initial decrease is due to the inclusion of many relatively small length on-times resulting from the micro-base stations in the voids of the homogeneous network. However, the volume fraction increases monotonically with the addition of micro-base stations.

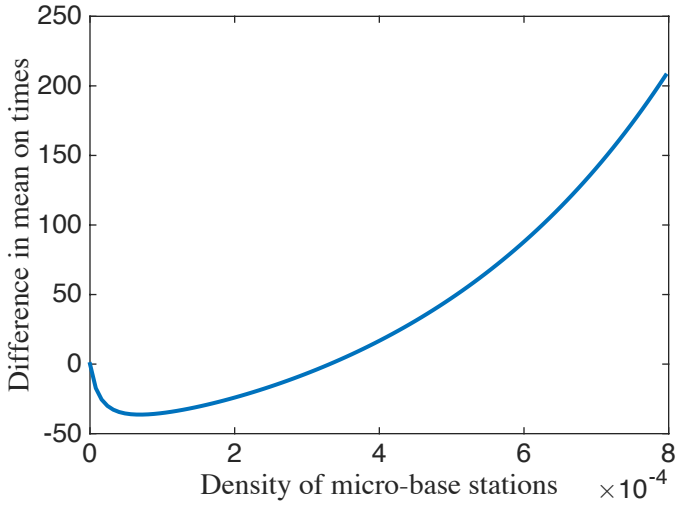


Figure 2.14: Comparing mean-on time for heterogeneous and homogeneous networks.

2.6.4 Stored Video Streaming in Heterogeneous Networks

Consider a scenario as discussed before where mobile users are viewing a video being streamed over a sequence of wireless downlinks, but now served by a heterogeneous network with micro and macro BS. In this setting once again we consider a fluid queue representing the tagged mobile user's playback buffer state similar to Section VII A, and ask whether there is a choice of γ such that the fluid queue is unstable, thus ensuring no rebuffering in the long term. Let the event Q denote that the tagged user is served, then

$$\rho(\gamma, \xi) = \frac{a \log(1+\gamma)\mathbb{P}(Q)}{\eta} \left(\mathbb{E}\left[\frac{1}{N_p^{(\gamma)} + 1} | Q \right] \right). \tag{2.31}$$

Assuming our approximation is valid, we need to find $E[\frac{1}{N_p^{(\gamma)}+1}]$ in the

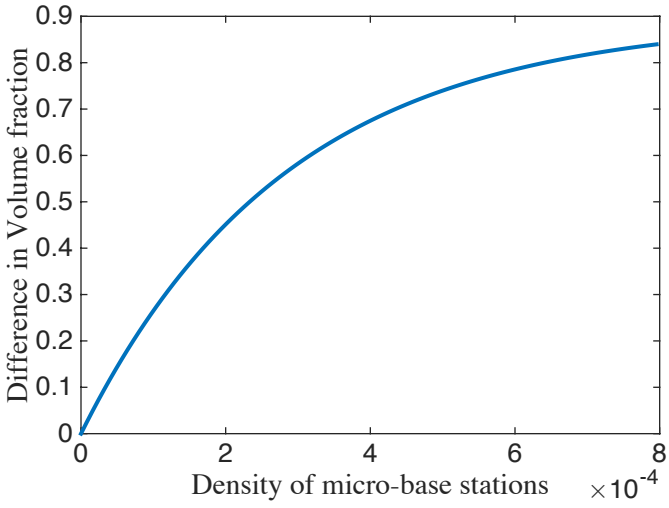


Figure 2.15: Comparing volume fraction for heterogeneous and homogeneous networks.

case of heterogeneous network. Let the events G and H be that the tagged user is served by micro BS and macro BS respectively.

$$\mathbb{E}\left[\frac{1}{N_p^{(\gamma)}+1}|Q\right] = \frac{\left(\mathbb{E}\left[\frac{1}{N_p^{(\gamma)}+1}|G\right]\mathbb{P}(G) + \mathbb{E}\left[\frac{1}{N_p^{(\gamma)}+1}|H\right]\mathbb{P}(H)\right)}{\mathbb{P}(Q)}. \tag{2.32}$$

Since, the micro and macro base stations are distributed independently, the mobile user experiences two independent alternating renewal processes. Thus, in the stationary regime, the probability that mobile is served by macro BS is the product of the probabilities that the mobile is "on" period of alternating renewal process of macro BS and in "off" period of that of micro BS

which gives:

$$\mathbb{P}(G) = 1 - e^{-\hat{\lambda}\pi\hat{r_\gamma}^2}, \mathbb{P}(H) = e^{-\hat{\lambda}\pi\hat{r_\gamma}^2}(1 - e^{-\lambda\pi r_\gamma^2}). \tag{2.33}$$

Let $\hat{J}_1^{(\gamma)}$ and $\hat{J}_2^{(\gamma)}$ denote the area similar to that of what we considered before. Given the mobile is served by macro BS, we need to consider the users which are within the area $\hat{J}_1^{(\gamma)}$ excluding the area covered by micro BS in this area. The area covered by the micro BS in $\hat{J}_1^{(\gamma)}$ is approximated to be $\nu^{(\hat{r}_\gamma)} \times \mathbb{E}[\hat{J}_1^{\gamma}]$, where $\nu^{(\hat{r}_\gamma)}$ is the fraction of space associated with micro-BS as given by (2.7). Then

$$\mathbb{E}\left[\frac{1}{N_p^{(\gamma)}+1}|G\right] = \frac{1 - e^{\xi\mathbb{E}[\hat{J}_2^{(\gamma)}]}}{\xi\mathbb{E}[\hat{J}_2^{(\gamma)}]},$$

$$\mathbb{E}\left[\frac{1}{N_p^{(\gamma)}+1}|H\right] = \frac{1 - e^{\xi\mathbb{E}[\hat{J}_1^{(\gamma)}](\hat{\lambda}\pi\hat{r}_\gamma^2)}}{\xi\mathbb{E}[\hat{J}_1^{(\gamma)}](\hat{\lambda}\pi\hat{r}_\gamma^2)}.$$

For a given density of macro-BS λ, the density of micro-BS $\hat{\lambda}$ and the density of users ξ, we can evaluate the SNR value γ^* for which the load factor $\rho(\gamma, \xi)$ given in (2.31) is maximum. Fig.2.16 illustrates the optimal SNR value (γ^*) with varying density of micro BS $\hat{\lambda}$. Notice that the optimal gamma value initially decreases with the density of micro-BS.

Since in our setting the micro BS have smaller transmission power, for a given threshold γ, the radius of coverage for micro BS is smaller than that of the macro BS .i.e., $\hat{r_\gamma} < r_\gamma$. Since the micro-BS are given higher preference, with an increase in their density, the optimal threshold decreases in order to increase their coverage. Also, the cost incurred by increasing coverage of

macro-BS is compensated by the increased density and coverage of micro-BS till a certain density.

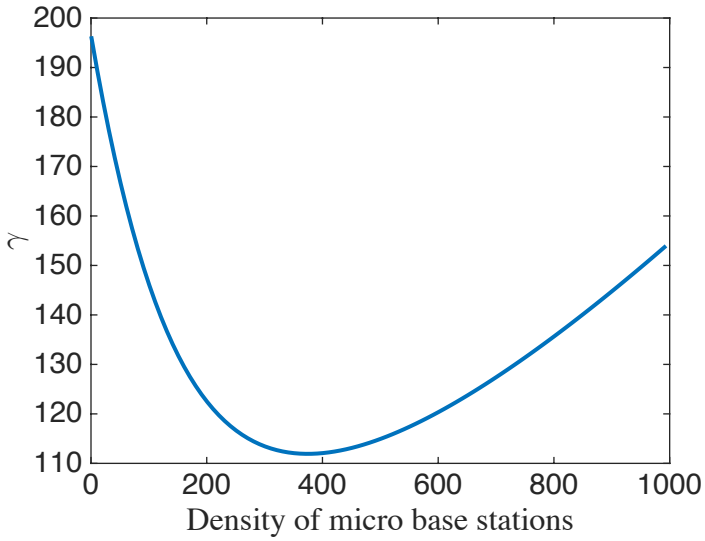


Figure 2.16: γ^* with increase in micro- BS density for a certain fixed density of macro-BS for an arbitrary positive constants a and η in heterogeneous case.

Fig.2.17 illustrates the level set curve of $\rho(\gamma^*, \xi) = 1$ for various values of $\hat{\lambda}$ and ξ for a given density of macro-BS λ. Given this setup, it is possible to answer questions like what is the minimum density of micro-base stations that needs to be deployed by the operator to serve a certain density of users, given the density of macro-BS λ such that the video streaming is uninterrupted for all the users. Thus, operators evaluate the cost incurred to serve a higher density of users by deploying micro-BS and in case the cost incurred is higher, an operator might consider other technologies.

Remark 2. *In heterogeneous networks, the interference from macro-BS to*

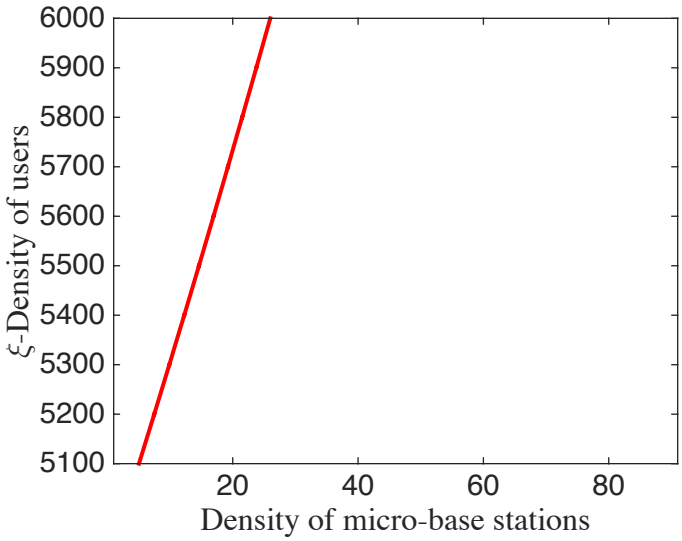


Figure 2.17: Level set curve of $\rho(\gamma^*, \xi) = 1$ for an arbitrary positive constants a and η in heterogeneous case.

micro-BS is of major concern. One can introduce blanking, to reduce the effect of such interference i.e., assume that with certain probability f micro-BS transmit and with probability $1 - f$ macro BS transmit. Also, by considering heterogeneous networks with cellular BS and WiFi hotspots, there is no such problem of interference since both cellular BS and WiFi hotspots operate at different frequencies.

2.7 Conclusion and Future Work

As explained in the introduction, the analysis of temporal variations of the shared rate experienced by a mobile user requires the characterization of both the functional distribution of a continuous parameter stochastic process

(rate process) constructed on a random spatial structure (e.g. the Poisson Voronoi tessellation) and of another stochastic process (sharing number process) constructed on a random distribution of static users. This chapter addressed the simplest question of this type by focusing on the underlying SNR process and the sharing number process in the absence of fading. This allowed us to derive an exact representation of the level crossings of the stochastic process of interest as an alternating renewal process with a full characterization of the involved distributions and of the asymptotic behavior of rare events. The simplicity and the closed form nature of this mathematical picture are probably the most important observations of the chapter. We also showed by simulation that this very simple model provides a good representation of the salient characteristics which happen for more complex systems, such as those with fading when the fading variance is small enough, or those based on SINR rather than SNR when the path loss exponent is large enough or those with changes in the direction of motion, when the change in direction is small. This model is hence of potential practical use as is, in addition to being a first glimpse at a set of new research questions.

The most challenging questions on the mathematical side are as follows: (1) The understanding of the tension between the randomness coming from geometry (studied in the present chapter), from sharing the network with other users (also studied in the present chapter) and that coming from propagation (only studied by simulation here): it would be nice to analytically quantify when one dominates the other. (2) The extension of the analysis to SINR

processes, which are our long term aim and will require significantly more sophisticated mathematical tools (e.g. based on functional distributions of shot noise fields), than those used so far.(3) The characterization of the distribution of the shared rate process. (4) The extension the case where all other users are mobile. To begin with, we might consider a stationary tagged user, with an independent homogeneous Poisson point process of mobile users moving on independent straight lines. Considering the Voronoi cell of the associated base station of the tagged user, we might try to characterize the number of mobile users entering and leaving the cell as an $M/GI/\infty$ queue. Thus, the sharing number process associated with the stationary user will be equivalent to the number of customers in a queue. Later, one could extend this to the case where the tagged user is mobile.

On the practical side, the main future challenges are linked to the initial motivations of this work, namely in the prediction and optimization of the user quality of experience. Many scenarios refining those studied can be considered. For instance, the stationary analysis of the fluid queue representing video streaming should be completed by a transient analysis and by a discrete time analysis. This alone opens an interesting and apparently unexplored connection between stochastic geometry and queuing theory with direct implications to video QoE.

Appendix

2.7.1 Busy Period of the M/GI/∞ Queue

Theorem 12. *(Makowski [21]) Consider an $M/GI/\infty$ queue with arrival rate λ and generic service time W. Let M denote an $\mathbb{N}$-valued random variable which is geometrically distributed according to*

$$\mathbb{P}(M = l) = (1-\nu)(\nu)^{l-1}, l = 1, 2, \ldots \tag{2.34}$$

with

$$\nu = 1 - e^{-\rho} \text{ and } \rho = \lambda\mathbb{E}[W]. \tag{2.35}$$

Consider the $\mathbb{R}^+$-valued random variable U distributed according to

$$\mathbb{P}(U \leq u) = \frac{1}{\nu}(1 - e^{-\rho\mathbb{P}(\hat{W} \leq u)}), u \geq 0, \tag{2.36}$$

where, $\hat{W}$ is the forward recurrence time associated with the generic service time W. Let $\{U_n, n \geq 1\}$ be an i.i.d. sequence independent of the random variable M. Let B denote a typical busy period. Then the forward recurrence time $\hat{B}$ associated with B admits the following random sum representation:

$$\hat{B} =_d \sum_{i=1}^{M} U_i, \tag{2.37}$$

where $=_d$ denotes equality in distribution.

2.7.2 Proof of Theorem 4

Lemma 1. *For a $\mathbb{R}^+$ -valued random variable $U^{(\gamma)}$ defined by (2.36) and service time distribution by (3.21). We have the following limit:*

$$\lim_{r_\gamma \to \infty} 2\lambda r_\gamma v\mathbb{E}[U^{(\gamma)}] = 1.$$

Proof. From Equations (2.7) and (2.36):

$$\mathbb{P}(U^{(\gamma)} > u) = \frac{e^{-\rho^{(\gamma)}}}{\nu^{(\gamma)}}(e^{\rho^{(\gamma)}\mathbb{P}(\hat{W}^{(\gamma)}>u)} - 1), u \geq 0, \tag{2.38}$$

and

$$\mathbb{E}[U^{(\gamma)}] = \int_0^{2r_\gamma/v} \mathbb{P}(U^{(\gamma)} > u)du = \frac{e^{-\rho^{(\gamma)}}}{\nu^{(\gamma)}} \int_0^{2r_\gamma/v} (e^{\rho^{(\gamma)}\mathbb{P}(\hat{W}^{(\gamma)}>u)} - 1)du. \tag{2.39}$$

Thus,

$$\mathbb{E}[U^{(\gamma)}] = \frac{e^{-\rho^{(\gamma)}}}{\nu^{(\gamma)}}\bigg(\int_0^{2r/v} e^{\rho^{(\gamma)}\mathbb{P}(\hat{W}^{(\gamma)}>u)}du - \frac{2r_\gamma}{v}\bigg). \tag{2.40}$$

Since $\hat{W}^{(\gamma)}$ is the forward recurrence time associated with the service time $W^{(\gamma)}$ defined in (2.6),

$$\mathbb{P}(\hat{W}^{(\gamma)} > u) = \frac{1}{\mathbb{E}[W^{(\gamma)}]} \int_u^\infty \mathbb{P}(W^{(\gamma)} > t)dt,$$

where, $\mathbb{E}[W^{(\gamma)}] = \frac{\rho^{(\gamma)}}{2r_\gamma v\lambda}$ from (2.7) and $\mathbb{P}(W^{(\gamma)} > t) = 1/2r_\gamma\sqrt{4r_\gamma^2 - v^2t^2}$ from (3.21). Thus,

$$\begin{aligned}
\mathbb{E}[U^{(\gamma)}] &= \frac{e^{-\rho^{(\gamma)}}}{\nu^{(\gamma)}}\bigg(\int_0^{\frac{2r_\gamma}{v}} e^{\lambda v\int_u^{\frac{2r_\gamma}{v}}\sqrt{4r_\gamma^2 - v^2t^2}dt}du - \frac{2r_\gamma}{v}\bigg) \\
&= \frac{e^{-\rho^{(\gamma)}}}{\nu^{(\gamma)}}\bigg(\int_0^{\frac{2r_\gamma}{v}} e^{\lambda\int_u^{2r_\gamma}\sqrt{4r_\gamma^2 - x^2}dx}du - \frac{2r_\gamma}{v}\bigg) \\
&= \frac{1}{\nu^{(\gamma)}}\bigg(\int_0^{\frac{2r_\gamma}{v}} e^{-\lambda\left(\frac{1}{2}u\sqrt{4r_\gamma^2 - v^2u^2} + 2r_\gamma^2\arctan(\frac{uv}{\sqrt{4r_\gamma^2 - v^2u^2}})\right)}du - e^{-\lambda\pi r_\gamma^2}\frac{2r_\gamma}{v}\bigg) \\
&= \frac{1}{1 - e^{-\lambda\pi r_\gamma^2}}\bigg(\frac{r_\gamma}{v}\int_0^2 e^{-\lambda r_\gamma^2 q(z)}dz - \frac{2e^{-\lambda\pi r_\gamma^2}r_\gamma}{v}\bigg).
\end{aligned} \tag{2.41}$$

The second equality is by the change of variables $vt = x$ and the last equality is by the change of variables $z = uv/r_\gamma$ and $q(z) = \frac{1}{2}z\sqrt{4-z^2} + 2\arctan(\frac{z}{\sqrt{4-z^2}})$. Now, from Equation (2.41) we have,

$$\lim_{r_\gamma \to \infty} \mathbb{E}[U^{(\gamma)}] \sim \frac{r_\gamma}{v} \int_0^2 e^{-\lambda r_\gamma^2 q(z)} dz.$$

Now, the integral $\frac{r_\gamma}{v}\int_0^2 e^{-\lambda r_\gamma^2 q(z)}$ and $\frac{r_\gamma}{v}\int_0^2 e^{-\lambda r_\gamma^2 2z} dz$ are asymptotically equal as r_γ goes to ∞.

The asymptotic equality of the two functions as r_γ goes to ∞ means, that the relative error of the approximate equality goes to 0 as r_γ goes to ∞ i.e.,

$$\lim_{r_\gamma \to \infty} \frac{\int_0^2 e^{-\lambda r_\gamma^2 q(z)} - e^{-\lambda r_\gamma^2 2z} dz}{\int_0^2 e^{-\lambda r_\gamma^2 2z}} = 0.$$

With the help of the Taylor series expansion, we get that

$$\lim_{r_\gamma \to \infty} \frac{\int_0^{C/r_\gamma^2} e^{-\lambda r_\gamma^2 q(z)} - e^{-\lambda r_\gamma^2 2z} dz}{\int_0^2 e^{-\lambda r_\gamma^2 2z}} = 0,$$

for some constant C. Thus, we need to prove that

$$\lim_{r_\gamma \to \infty} \frac{\int_{C/r_\gamma^2}^2 e^{-\lambda r_\gamma^2 q(z)} - e^{-\lambda r_\gamma^2 2z} dz}{\int_0^2 e^{-\lambda r_\gamma^2 2z}} = 0.$$

Now let us observe the function $h(z) = e^{-\lambda r_\gamma^2 q(z)} - e^{-\lambda r_\gamma^2 2z}$. The derivative of the function $h(z)$ is

$$\begin{aligned} h'(z) &= \big(-\sqrt{4-z^2} e^{-\lambda r_\gamma^2 q(z)} + 2e^{-\lambda r_\gamma^2 2z} \big) r_\gamma^2 \\ &= e^{-\lambda r_\gamma^2 2z} r_\gamma^2 \left[2 - \sqrt{4-z^2} e^{-\lambda r^2 [q(z) - 2z]} \right]. \end{aligned} \tag{2.42}$$

Thus by substituting $z = C/r_\gamma^2$ and using the Taylor series expansion of $[q(z) - 2z]$ and $\sqrt{4 - z^2}$, we get

$$\lim_{r_\gamma \to \infty} h'(C/r_\gamma^2) = e^{-2C\lambda z}[o(r_\gamma)] = 0.$$

The derivative of function $h(z)$ is 0 for $z = C/r_\gamma^2$ as r_γ goes to ∞ i.e., the function $h(z)$ has a local maximum at C/r_γ^2 for some constant C and it can be observed that it is decreasing for $C = 3$.

Therefore, the function $h(z)$ is a decreasing function from $3/r_\gamma^2$ to 2 and the area under its curve has an upper bound that goes to 0. Thus, the two integrals are asymptotically equal as r_γ goes to ∞, which implies that $\mathbb{E}[U^{(\gamma)}] \sim \frac{1}{2vr_\gamma\lambda}$ as r_γ goes to ∞. □

Proof. $V^{(\gamma)}$ is the sum of a busy period $B^{(\gamma)}$ and idle period $I^{(\gamma)}$ of an $M/GI/\infty$ queue. We know that the distribution of the idle period is exponential with parameter $2\lambda v r_\gamma$ and that the busy period $B^{(\gamma)}$ and idle period $I^{(\gamma)}$ are independent.

The Laplace transform of the random variable $V^{(\gamma)}$ scaled by $2\lambda v r_\gamma$ is given by

$$\Phi_{V^{(\gamma)}}(f(\gamma)s) = \Phi_{B^{(\gamma)}}(f(\gamma)s)\Phi_{I^{(\gamma)}}(f(\gamma)s), \tag{2.43}$$

where,

$$\Phi_{I^{(\gamma)}}(f(\gamma)s) = \frac{1}{1+s}. \tag{2.44}$$

So, asymptotically, as r_γ goes to 0 , $\Phi_{I^{(\gamma)}}(f(\gamma)s)$ converges in distribution to an exponential random variable with unit mean.

Next, we need to prove that asymptotically, the busy period $B^{(\gamma)}$ scaled by $f(\gamma)$ goes to one. Let us consider the forward recurrence time of the busy period, $\hat{B}^{(\gamma)}$ which is defined in (2.6). From Theorem 2 we get,

$$\mathbb{E}[e^{-s\hat{B}^{(\gamma)}}] = \mathbb{E}[e^{-s\sum_{i=1}^{M^{(\gamma)}} U_i^{(\gamma)}}] = \frac{(1-\nu^{(\gamma)})\mathbb{E}[e^{-sU^{(\gamma)}}]}{1-\nu^{(\gamma)}\mathbb{E}[e^{-sU^{(\gamma)}}]},$$

and it follows that

$$\mathbb{E}[e^{-sf(\gamma)\hat{B}^{(\gamma)}}] = \frac{(1-\nu^{(\gamma)})\mathbb{E}[e^{-sf(\gamma)U^{(\gamma)}}]}{1-\nu^{(\gamma)}\mathbb{E}[e^{-sf(\gamma)U^{(\gamma)}}]}. \tag{2.45}$$

Further, we have that

$$\mathbb{E}[e^{-sf(\gamma)U^{(\gamma)}}] = \sum_{k=0}^{\infty} l_k(s, r_\gamma), \tag{2.46}$$

where

$$l_k(s, r_\gamma) = \frac{(-1)^k (2\lambda v s r_\gamma)^k \mathbb{E}[(U^{(\gamma)})^k]}{k!}.$$

Given that the support of service times $W^{(\gamma)}$ is $[0, 2r_\gamma/v]$, it follows from (2.36) that the support of the random variable $U^{(\gamma)}$ is also $[0, 2r_\gamma/v]$ and all its moments are bounded $\mathbb{E}[(U^{(\gamma)})^k] < (2r_\gamma/v)^k$. Thus, for all values of s, we get that $\lim_{r_\gamma \to 0} l_k(s, r_\gamma) = 0$. For all $k = 2, 3..$ we also have that

$$|l_k(s, r_\gamma)| \leq \frac{(2\lambda v s r_\gamma^2)^k}{k!}.$$

Therefore, for all values of s, $\lim_{r_\gamma \to 0} \sum_{k=2}^{\infty} l_k(s, r_\gamma) = 0$. Thus, Eq (2.46) can be written as

$$\mathbb{E}[e^{-sf(\gamma)U^{(\gamma)}}] = 1 - s2\lambda r_\gamma v \mathbb{E}[U^{(\gamma)}] + o(r_\gamma).$$

Now, substituting into Eq (2.45) we have

$$\mathbb{E}[e^{-sf(\gamma)\hat{B}^{(\gamma)}}] = \frac{1-\nu^{(\gamma)}+2s\lambda v r_\gamma \mathbb{E}[U^{(\gamma)}](1-\nu^{(\gamma)})+o(r_\gamma)}{1-\nu^{(\gamma)}+2\nu^{(\gamma)}s\lambda v r_\gamma \mathbb{E}[U^{(\gamma)}]+o(r_\gamma)}, \tag{2.47}$$

from which it follows that

$$\lim_{r_\gamma \to 0} \mathbb{E}[e^{-sf(\gamma)\hat{B}^{(\gamma)}}] = 1. \tag{2.48}$$

From the results in [21] we have that

$$\mathbb{E}[e^{-sf(\gamma)B^{(\gamma)}}] = 1 - sf(\gamma)\mathbb{E}[B^{(\gamma)}]\mathbb{E}[e^{-sf(\gamma)\hat{B}^{(\gamma)}}], \tag{2.49}$$

and also that

$$\mathbb{E}[B^{(\gamma)}] = \frac{\nu^{(\gamma)}}{2\lambda v r_\gamma (1-\nu^{(\gamma)})}. \tag{2.50}$$

Thus, we get

$$\mathbb{E}[e^{-sf(\gamma)B^{(\gamma)}}] = 1 - s\frac{1-e^{-\lambda\pi r_\gamma^2}}{e^{-\lambda\pi r_\gamma^2}}\mathbb{E}[e^{-sf(\gamma)\hat{B}^{(\gamma)}}].$$

Using the limit in (2.48), we get

$$\lim_{r_\gamma \to 0} \mathbb{E}[e^{-sf(\gamma)B^{(\gamma)}}] = 1. \tag{2.51}$$

Thus, the distribution of the random variable $V^{(\gamma)}$ scaled by $f(\gamma) = 2\lambda v r_\gamma$ converges in distribution to an exponential random variable with unit mean.

Now, consider a continuous function $g(\gamma)$ such that $\lim_{r_\gamma \to \infty} g(\gamma) = 0$. The Laplace transform of the random variable $V^{(\gamma)}$ scaled by $g(\gamma)$ is given by

$$\Phi_{V^{(\gamma)}(s)} = \Phi_{B^{(\gamma)}}(g(\gamma)s)\Phi_{I^{(\gamma)}}(g(\gamma)s), \tag{2.52}$$

where,

$$\Phi_{I^{(\gamma)}}(g(\gamma)s) = \frac{1}{1 + sg(\gamma)(1/2vr_\gamma\lambda)}. \tag{2.53}$$

Asymptotically, as r_γ goes to ∞, $\Phi_{I^{(\gamma)}}(g(\gamma)s)$ goes to 1. By the same arguments as those for the up-crossing case we have that

$$\mathbb{E}\left[e^{-sg(\gamma)\hat{B}^{(\gamma)}}\right] = \frac{(1-\nu^{(\gamma)})\mathbb{E}\left[e^{-sg(\gamma)U^{(\gamma)}}\right]}{1-\nu^{(\gamma)}\mathbb{E}\left[e^{-sg(\gamma)U^{(\gamma)}}\right]},$$

where,

$$\mathbb{E}\left[e^{-sg(\gamma)U^{(\gamma)}}\right] = \sum_{k=0}^{\infty}\frac{(-1)^k(2\lambda vsr_\gamma)^k(e^{-\rho^{(\gamma)}})^k}{k!}\mathbb{E}\left[(U^{(\gamma)})^k\right]. \tag{2.54}$$

Let $m_k(s, r_\gamma) = \frac{(-1^k)(2\lambda vsr_\gamma)^k(e^{-\rho^{(\gamma)}})^{k-1}}{k!}\mathbb{E}[(U^{(\gamma)})^k]$. The moments of the random variable $U^{(\gamma)}$ are bounded $\mathbb{E}[(U^{(\gamma)})^k] < (2r_\gamma/v)^k$. Thus, for all values of s, we get that $\lim_{r_\gamma\to\infty} m_k(s, r_\gamma) = 0$, because $\rho^{(\gamma)} = \lambda\pi r_\gamma^2$ and the exponential term $(e^{-\rho^{(\gamma)}})^{k-1}$ dominates. Also for $k = 2, 3, ...$ we have that

$$|m_k(s, r_\gamma)| \leq \frac{(2\lambda vsr_\gamma^2)^k(e^{-\rho^{(\gamma)}})^{k-1}}{k!}.$$

Therefore for all values of s,

$$\lim_{r_\gamma\to\infty}\sum_{k=2}^{\infty} m_k(s, r_\gamma) = 0.$$

Thus, Eq (2.54) can be written as

$$\mathbb{E}[e^{-sf(\gamma)U^{(\gamma)}}] = 1 - se^{-\rho^{(\gamma)}}2vr_\gamma\lambda\mathbb{E}[U^{(\gamma)}] + e^{-\rho^{(\gamma)}}o(r_\gamma),$$

where, $\lim_{r_\gamma \to \infty} o(r_\gamma) = 0$.

From Lemma 1 we have that, $\lim_{r_\gamma \to \infty} 2\lambda r_\gamma v \mathbb{E}[U^{(\gamma)}] = 1$. Putting these results together we have that

$$\mathbb{E}[e^{-sg(\gamma)\hat{B}^{(\gamma)}}] = \frac{e^{-\rho^{(\gamma)}} - se^{-2\rho^{(\gamma)}} 2vr_\gamma \lambda \mathbb{E}[U^{(\gamma)}] + e^{-2\rho^{(\gamma)}} o(r_\gamma)}{e^{-\rho^{(\gamma)}}\left[1 - e^{-\rho^{(\gamma)}}\right]\left[s2vr_\gamma \lambda \mathbb{E}[U^{(\gamma)}] - e^{-\rho^{(\gamma)}} o(r_\gamma)\right] + e^{-\rho^{(\gamma)}}}, \tag{2.55}$$

so that

$$\lim_{r_\gamma \to \infty} \mathbb{E}[e^{-sg(\gamma)\hat{B}^{(\gamma)}}] = \frac{1}{1+s}. \tag{2.56}$$

Now, from Eq (2.49) and (2.50) we get,

$$\mathbb{E}[e^{-sf(\gamma)B^{(\gamma)}}] = 1 - s(1 - e^{-\lambda\pi r_\gamma^2})\mathbb{E}[e^{-sf(\gamma)\hat{B}^{(\gamma)}}].$$

Thus from the limit in (2.56), $\lim_{r_\gamma \to \infty} \mathbb{E}[e^{-sg(\gamma)B^{(\gamma)}}] = \frac{1}{1+s}$.

Thus, the random variable $V^{(\gamma)}$ scaled by $g(\gamma) = 2\lambda v r_\gamma e^{-\lambda\pi r_\gamma^2}$ converges in distribution to an exponential random variable with unit mean.

□

2.7.3 Proof of Theorem 6

Let us first prove the second relation. Let $K = p/w$ and let D denote the distance to the closest base station. We have

$$\begin{aligned}
\mathbb{P}((\log(1+\mathrm{SNR}) > s) &= \mathbb{P}\left(KD^{-\beta} > e^s - 1\right) = \mathbb{P}\left(D^\beta < \frac{K}{e^s - 1}\right) \\
&= \mathbb{P}\left(D^2 < \left(\frac{K}{e^s - 1}\right)^{\frac{2}{\beta}}\right) = 1 - \exp\left(-\lambda\pi\left(\frac{K}{e^s - 1}\right)^{\frac{2}{\beta}}\right),
\end{aligned}$$

where we used the fact that D^2 is an exponential random variable with parameter $\lambda\pi$. The result then follows from the bound $a \geq 1 - e^{-a} \geq a + a^2/2$, for $a \geq 0$.

We now prove the first relation. Since $\hat{S}^{(\gamma)} \leq S^{(\gamma)} \leq \log(1 + \mathrm{SNR})$, in order to prove the first inequality, it is enough to show that

$$\lim_{s\to\infty} -\frac{1}{s} \log\left(\mathbb{P}(\hat{S}^{(\gamma)} > s)\right) \leq \frac{2}{\beta}. \tag{2.57}$$

We have

$$\mathbb{P}(\hat{S}^{(\gamma)} > s) \;=\; \mathbb{P}(\log(1 + KD^{-\beta}) > s(1 + \hat{N}_\gamma)),$$

hence

$$\begin{aligned}
\mathbb{P}(\hat{S}^{(\gamma)} > s) &= \mathbb{P}\left(KD^{-\beta} > \exp\left(s(1+\hat{N}_\gamma)\right) - 1\right) \\
&\geq \mathbb{P}\left(KD^{-\beta} > \exp\left(s(1+\hat{N}_\gamma)\right)\right) \\
&= \mathbb{P}\left(D^\beta < K\exp\left(-s(1+\hat{N}_\gamma)\right)\right) \\
&= \mathbb{P}\left(D^2 < K^{\frac{2}{\beta}}\exp\left(-s\frac{2}{\beta}(1+\hat{N}_\gamma)\right)\right) \\
&= \mathbb{P}\left(D^2 < K^{\frac{2}{\beta}} e^{-s\frac{2}{\beta}} e^{-s\frac{2}{\beta}\hat{N}_\gamma}\right).
\end{aligned}$$

Using now the fact that D^2 is an exponential random variable with parameter $\lambda\pi$, independent of $\hat{N}_\gamma$, we get

$$\begin{aligned}
\mathbb{P}(\hat{S}^{(\gamma)} > s) &\geq 1 - \mathbb{E}\left(\exp\left(-\lambda\pi K^{\frac{2}{\beta}} e^{-s\frac{2}{\beta}} e^{-s\frac{2}{\beta}\hat{N}_\gamma}\right)\right) \\
&\geq \lambda\pi K^{\frac{2}{\beta}} e^{-s\frac{2}{\beta}} \exp\left(-\xi\pi r_\gamma^2\left(1 - e^{-s\frac{2}{\beta}}\right)\right)\left(1 + o\left(e^{-s\frac{2}{\beta}}\right)\right),
\end{aligned}$$

where we used again the bound $1 - e^{-a} \geq a + a^2/2$ and the fact that $\hat{N}^{(\gamma)}$ is Poisson with parameter $\xi\pi r_\gamma^2$.

Taking the log, multiplying both sides by $-\frac{1}{s}$ and letting s go to infinity gives (2.57).

2.7.4 Proof of Theorem 7

Let us first prove the last relation. Using the Chernoff's bound for the Poisson random variable $\hat{N}^{(\gamma)}$, we get

$$\mathbb{P}\left(\hat{N}^{(\gamma)}+1>s_n\log(1+Kr_\gamma^{-\beta})\right)\leq e^{-\theta}\left(\frac{e\theta}{s_n\log(1+Kr_\gamma^{-\beta})}\right)^{s\log(1+Kr_\gamma^{-\beta})},$$

with $\theta=\xi\pi r_\gamma^2$, which shows that

$$\lim_{n\to\infty}-\frac{1}{s_n\log(s_n)}\log\left(\mathbb{P}\left(\hat{N}^{(\gamma)}+1>s_n\log(1+Kr_\gamma^{-\beta})\right)\right)\geq\log(1+Kr_\gamma^{-\beta}).$$

The converse inequality follows from the bound

$$\mathbb{P}\left(\hat{N}^{(\gamma)}+1>s_n\log(1+Kr_\gamma^{-\beta})\right)\geq e^{-\theta}\frac{\theta^k}{k!},$$

with $k=s_n\log(1+Kr_\gamma^{-\beta})$ and Stirling's bound on the Gamma function.

Let us now prove the first relation. We have

$$\mathbb{P}\left(\hat{S}^{(\gamma)}<\frac{1}{s_n}\right)=\int_0^{r_\gamma^2}e^{-\lambda\pi v}\mathbb{P}\left(\hat{N}^{(\gamma)}+1>s_n\log\left(1+Kv^{-\frac{\beta}{2}}\right)\right)\mathrm{d}v.$$

For s_n large enough,

$$\mathbb{P}\left(\hat{S}^{(\gamma)}<\frac{1}{s_n}\right)\leq\mathbb{P}\left(\hat{N}^{(\gamma)}+1>s_n\log\left(1+Kr_\gamma^{-\frac{\beta}{2}}\right)\right)\int_0^{r_\gamma^2}e^{-\lambda\pi v}\mathrm{d}v,$$

which allows one to conclude that

$$\lim_{n\to\infty}-\frac{1}{s_n\log(s_n)}\log\left(\mathbb{P}\left(\hat{S}^{(\gamma)}<\frac{1}{s_n}\right)\right)\geq\log(1+Kr_\gamma^{-\beta}).$$

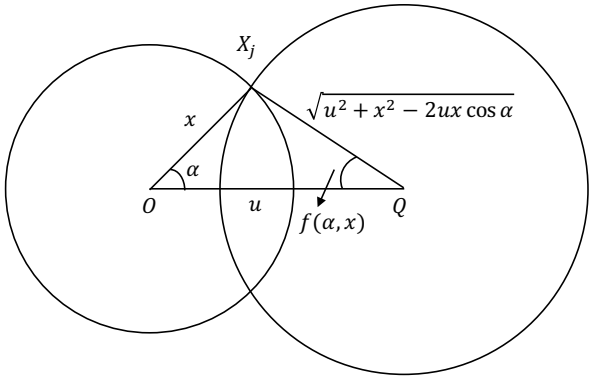


Figure 2.18: Tagged user at origin, its serving base station X_j at distance x and a point Q at a distance u along with discs $B_x(0)$ and $B_z(Q)$.

The upper bound follows from the inequality:

$$\begin{aligned}\mathbb{P}\left(\hat{S}^{(\gamma)} < \frac{1}{s_n}\right) &= \sum_{l=0}^{\infty} e^{-\theta}\frac{\theta^l}{l!}\mathbb{P}\left(D^2 > \left(\frac{K}{e^{\frac{l+1}{s}}-1}\right)^{\frac{2}{\beta}}\right) \\ &\geq e^{-\theta}\frac{\theta^k}{k!}\mathbb{P}\left(D^2 > \left(\frac{K}{e^{\frac{k+1}{s}}-1}\right)^{\frac{2}{\beta}}\right),\end{aligned}$$

where $k = s_n \log(1 + Kr_\gamma^{-\beta})$.

2.7.5 Integral Geometry I

Consider a tagged user at the origin, and let the serving BS X_j be at a distance x from the origin. Let Q denote a point a distance u from the origin. We need to consider all points Q that are within a distance r from the serving BS X_j i.e., $z = |X_jQ| = \sqrt{u^2+x^2-2ux\cos(\alpha)} < r$, where $\alpha = \angle X_jOQ$ as shown in Fig.8.

Let $B_x(0), B_z(Q)$ be two discs with centers at the origin and Q, and

radii x and X_jQ, respectively.

The conditional probability that a point at distance u from origin is within the Voronoi cell of the BS serving the user at the origin given it is at a distance x is the probability that there is no other BS in the area of disc $B_z(Q)$ excluding the area of $B_x(0)$. Then the conditional expectation is:

$$\mathbb{E}[V^*|D=x] = 2\int_0^{\pi}\int_0^{x\cos(\alpha)+\sqrt{r^2-x^2\sin(\alpha)}} e^{(-\lambda l(B_z(Q)-B_x(0)))}ududu\alpha, \quad (2.58)$$

where, D is a random variable denoting the distance to the closest BS. Thus,

$$\begin{aligned}\mathbb{E}[V^*] &= \int_0^r \mathbb{E}[V^*|D=x]\frac{f_D(x)}{\int_0^r f_D(y)dy}dx = \frac{\int_0^r \mathbb{E}[V^*|D=x]2\lambda\pi x e^{-\lambda\pi x^2}dx}{1-e^{-\lambda\pi r^2}} \\ &= \frac{1}{1-e^{-\lambda\pi r^2}}\int_0^r 2\int_0^{\pi}\int_0^{x\cos(\alpha)+\sqrt{r^2-x^2\sin(\alpha)}} e^{(-\lambda l(B_z(Q)-B_x(0)))}ududu\alpha 2\lambda\pi x e^{-\lambda\pi x^2}dx \\ &= \frac{1}{1-e^{-\lambda\pi r^2}}4\lambda\pi\int_0^r\int_0^{\pi}\int_0^{x\cos(\alpha)+\sqrt{r^2-x^2\sin(\alpha)}} e^{(-\lambda l(B_x(0)\cup B_z(Q)))}uxdud\alpha dx. \end{aligned} \quad (2.59)$$

The last equality is because $l(B_x(0)) = \pi r^2$ and $l(B_x(0)\cup B_z(Q)) = ux\sin(\alpha)+(\pi-\alpha)x^2+(\pi-f(\alpha,x))(u^2+x^2-2ux\cos(\alpha))$ with $\cos(f(\alpha,x)) = \frac{u-x\cos(\alpha)}{\sqrt{u^2+x^2-2ux\cos(\alpha)}}$.

2.7.6 Integral Geometry II

Let us consider a point U at a distance u from the origin on the line $d(0,\theta)$ as shown in the figure. Then, the probability that the point is in the Voronoi cell of the BS X_j is given by $e^{-\lambda l(C(u,\theta))}$, where $l(C(u,\theta))$ is the area

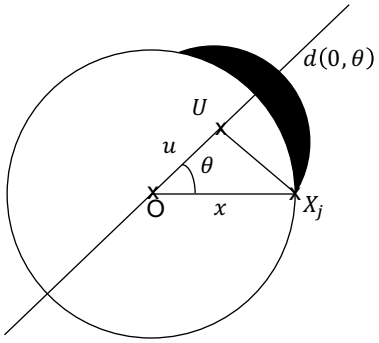


Figure 2.19: The tagged user at the origin, its serving base station X_j at distance x and a point U at a distance u on the line $d(0,\theta)$.

of the shaded region in the figure. The conditional expectation is given by:

$$\mathbb{E}[l(0,\theta)|D=x] = \frac{1}{\pi}\int_0^{\pi}\int_0^{x\cos(\theta)+\sqrt{r^2-x^2\sin(\theta)}} e^{(-\lambda l(C(u,\theta)))}dud\theta, \tag{2.60}$$

where, D is the random variable denoting the distance to the closest BS.

Let $B_x(0), B_z(U)$ be two discs with centers at the origin and U, and radii x and X_jU, respectively. Thus,

$$\begin{aligned}\mathbb{E}[l(0,\theta)] &= \int_0^r \mathbb{E}[l(0,\theta)|D=x]\frac{f_R(x)}{\int_0^r f_D(y)dy}dx = \frac{\int_0^r \mathbb{E}[l(0,\theta)|D=x]2\lambda\pi x e^{-\lambda\pi x^2}dx}{1-e^{-\lambda\pi r^2}} \\ &= \frac{1}{\pi(1-e^{-\lambda\pi r^2})}\int_0^r\int_0^{\pi}\int_0^{x\cos(\theta)+\sqrt{r^2-x^2\sin(\theta)}} e^{(-\lambda l(C(u,\theta)))}dud\theta 2\lambda\pi x e^{-\lambda\pi x^2}dx \\ &= \frac{1}{\pi(1-e^{-\lambda\pi r^2})}2\lambda\pi\int_0^r\int_0^{\pi}\int_0^{x\cos(\theta)+\sqrt{r^2-x^2\sin(\theta)}} e^{(-\lambda l(B_x(0)\cup B_z(U)))}xdud\theta dx.\end{aligned} \tag{2.61}$$

The last equality is because $l(B_x(0)) = \pi r^2$ and $l(B_x(0)\cup B_z(U)) = ux\sin(\theta)+(\pi-\theta)x^2+(\pi-f(\theta,x))(u^2+x^2-2ux\cos(\theta))$ with $\cos(f(\theta,x)) = \frac{u-x\cos(\theta)}{\sqrt{u^2+x^2-2ux\cos(\theta)}}$.

2.7.7 Proof of Theorem 9

Proof. Let Z be a geometric random variable given by:

$$P(Z = l) = (1 - \alpha)\alpha^l. \tag{2.62}$$

Then, the random variable T representing the time taken to download the file is the geometric sum of independent random variables given by

$$T = \sum_{i=0}^{Z} (X_i + I_i^{(\gamma)}) + Y. \tag{2.63}$$

Thus, the Laplace transform of T is

$$\begin{aligned} \mathcal{L}_T(s) &= \left[(1-\alpha) + (1-\alpha)\alpha\mathcal{L}_X(s)\mathcal{L}_{I^{(\gamma)}}(s) + ..\right]\mathcal{L}_Y(s) \\ &= \frac{(1-\alpha)\mathcal{L}_Y(s)}{1 - \alpha\mathcal{L}_X(s)\mathcal{L}_{I^{(\gamma)}}(s)} = \frac{(1-\alpha)\mathcal{L}_Y(s)}{1 - \alpha\mathcal{L}_X(s)\frac{2\lambda vr}{2\lambda vr+s}}. \end{aligned} \tag{2.64}$$

The last equality follows from the fact that $I^{(\gamma)}$ is exponential with parameter $2\lambda vr$. □

Chapter 3

Shannon Rate Fields in Interference Limited Environments

3.1 Introduction and Related Work

The aim in this chapter is to study the characteristics of spatial fields associated with ultra dense wireless networks and leverage them to study the temporal characteristics of mobile users. We also link them to basic engineering questions such as: (1) What are the variations in interference over a time period, since such variability can be a challenge to techniques such as adaptive modulation and coding which rely on the predictability of SINR over time periods, and (2) how the backhaul capacity requirement per base station scales with densification?

We focus on studying scaling limits for ultra dense networks for more

Portions of this chapter has been published as: Madadi, Pranav, Franois Baccelli, and Gustavo de Veciana. "On Spatial and Temporal Variations in Ultra Dense Wireless Networks." In IEEE INFOCOM 2018-IEEE Conference on Computer Communications, pp. 2663-2671. IEEE, 2018.
Co-authors have participated extensively in model formulation and research methods, and have contributed in reviewing the final manuscript.

realistic models. With the increases in density, the distance between the users and their associated base stations reduces leading to an increase in the signal strength, but there is also an increase in interference. Together this effectively reduces the impact of thermal noise. Thus, ultra dense networks can be interference limited networks[1]. Further, for these short range inter-site distances, the unbounded path loss models often used in the literature are clearly no longer physically relevant as they are singular at the origin. More realistic and practical models would be based on bounded path loss functions.

Related Work. Cellular network performance has been extensively studied by modeling the network using stochastic geometry [32], [6]. With the help of scaling limits for the interference, network performance has been evaluated under densification for various modeling assumptions ([33], [34], [35], [36], [37], [38]). Coverage probability and area spectral efficiency analysis have mostly been used as the main performance metrics. The findings are sometimes conflicting and suggest that densification may eventually stop delivering significant throughput gains.

Most prior work focuses on either studying the scaling limits of the SINR at a *typical location*, or on two-point correlations in interference and shadowing. [39], [40], [41]. Although, [42] studies the scaling limit of the interference field with singular power law path loss and Rayleigh fading, to

[1] An argument can be made that beamforming can mitigate this in part. Here we will assume such narrow band beamforming, if present is still broad and focus on omni-directional transmissions for simplicity.

our knowledge, a spatial characterization of the limiting interference field for bounded path loss models is lacking.

There is a broad body of relevant work in the field of mathematics of shot noise and Gaussian fields. These are of interest to this class of problems since the interference fields in large wireless systems can be viewed as shot noise fields, where the path loss function is equivalent to the kernel function of the shot noise field. Further, the infinite divisibility property of Poisson point processes allows one to establish convergence of the shot noise field to a Gaussian field as the intensity increases ([43] [44]). Precise sample path properties, especially level crossings of shot noise fields have been extensively studied in, e.g., [45], [46], [47], [48]. General results are known concerning the level crossings of smooth Gaussian processes ([45], [49]). However, with closest base station association, the interference is not a shot noise field but a protected shot noise field. Thus, most of the results in the literature are not directly applicable.

Given the importance of backhauling for 5G small cell networks, many researchers have studied centralized and distributed architectures for the backhauling gateways, see, e.g., [50]. Simulation results suggest that a distributed wireless backhaul network architecture is more suitable for future 5G networks employing massive MIMO and/or millimeter wave communication technologies. Millimeter wave communication has been considered as the wireless backhaul solution for small cell networks in 5G communication systems. However, most studies on millimeter wave backhaul technologies focus on the design of

the antenna array and radio frequency (RF) components of transceivers, such as beamforming and modulation schemes [51], [52]. To our knowledge, a system level investigation of ultra-dense cellular networks backhaul requirements such as that in this paper is novel.

Adaptive modulation and coding is a critical technique for adapting to time varying channels resulting from, say fading, path loss and interference, see, e.g., [53], [54] and the references therein. Most of the work focuses on channel quality estimation or various adaptive modulation techniques. In this paper, we characterize the time periods over which we can predict SINR with some success and study their dependence on various system parameters.

3.2 System Model

Consider a cellular network where the base stations are distributed according to a homogeneous Poisson Point Process (PPP) $\Phi = \{\mathbf{X}_1, \mathbf{X}_2, ..\}$ in $\mathbb{R}^2$ of intensity λ_b. Let $l : \mathbb{R}^2 \to \mathbb{R}_+$ be a deterministic non-negative function. Consider downlink transmissions and assume all base stations transmit at a fixed power p. Then, the total power received from all base stations at a location $\mathbf{y}$ is referred to as a shot noise field [55], given by $\mathcal{I}_{\lambda_b}(\mathbf{y}) = \sum_{\mathbf{X}_i \in \Phi} pl(\mathbf{X}_i - \mathbf{y})$, where $\mathbf{y} \in \mathbb{R}^2$.

Assuming that a user associates with the closest base station, the total interference seen by the user is given by the protected shot noise field:

$$\mathcal{J}_{\lambda_b}(\mathbf{y}) = \sum_{\mathbf{X}_i \in \Phi \setminus \mathbf{X}_{\lambda_b}(\mathbf{y})} pl(\mathbf{X}_i - \mathbf{y}), \tag{3.1}$$

where, $\mathbf{X}_{\lambda_b}(\mathbf{y}) \in \Phi$ denotes the closest base station in Φ to location $\mathbf{y}$. Given the interference field, one can study the SIR field and Shannon rate field experienced by a user at a given location. In the absence of fading and shadowing, the SIR field $(\mathrm{SIR}_{\lambda_b}(\mathbf{y}), \mathbf{y} \in \mathbb{R}^2)$ is given by

$$\mathrm{SIR}_{\lambda_b}(\mathbf{y}) = \frac{p * l(\mathbf{y} - \mathbf{X}_{\lambda_b}(\mathbf{y}))}{\mathcal{J}_{\lambda_b}(\mathbf{y})}. \tag{3.2}$$

In the interference limited regime, the Shannon rate field, $(\mathcal{S}_{\lambda_b}(\mathbf{y}), \mathbf{y} \in \mathbb{R}^2)$ can be defined from the SIR field through

$$\mathcal{S}_{\lambda_b}(\mathbf{y}) = w \log(1 + \mathrm{SIR}_{\lambda_b}(\mathbf{y})), \tag{3.3}$$

where, w is the wireless system bandwidth.

In the sequel we will also consider a tagged user moving at a fixed unit velocity along a straight line starting from the origin at time $t = 0$. We will denote the interference and the Shannon rate experienced by the mobile user by the stochastic processes $(J_{\lambda_b}(t), t > 0)$ and $(S_{\lambda_b}(t), t > 0)$, respectively. Note that $\mathcal{J}_{\lambda_b}$ and $\mathcal{I}_{\lambda_b}$ correspond to spatial processes or fields while J_{λ_b} and I_{λ_b} are temporal processes. Table 3.1 summarizes the notation used in the chapter.

In order to explore the limiting characteristics of the random fields associated with dense networks, we will need to properly define weak convergence notations for the two-dimensional processes. Recall that the D_1 functional space is the set of functions defined on reals that are right continuous with left limits, also referred to as cadlag functions. Note that the space of continuous

functions is in D_1. Similarly, if T is the unit square $[0,1]^2$, then the functional space D_2 is the uniform closure, in the space of all bounded functions from T to $\mathbb{R}$, of the vector subspace of simple functions which are coordinate wise D_1 [56]. The reason for this choice is that weak convergence of stochastic processes can be studied in this functional space under the S-topology [56]. We further classify the path loss functions in the D_2 functional space as follows:

1. Class-1 Functions : functions that are smooth, integrable and twice differentiable i.e., $l \in C^2(\mathbb{R}^2), l \in L^2(\mathbb{R}^2)$, $l^{'}, l^{"} \in L^1(\mathbb{R}^2)$. The stretched exponential path loss function, $l_1(\mathbf{y}) = e^{-a_1||\mathbf{y}||^4}$, [57] for a constant a_1, is an example of Class-1 function.

2. Class-2 Functions : functions that are continuous, piece-wise C^2 on $\mathbb{R}^2$, with discontinuities in their first derivative. Multi-slope path loss functions are examples of Class-2 [35]:

$$l_2(\mathbf{y}) = \begin{cases} 1 & \text{for } ||\mathbf{y}|| \leq r_0, \\ a_1/||\mathbf{y}||^{\beta_1} & \text{for } r_0 < ||\mathbf{y}|| \leq r_1, \\ a_2/||\mathbf{y}||^{\beta_2} & \text{for } ||\mathbf{y}|| > r_1. \end{cases} \tag{3.4}$$

   where, a_1 and a_2 are constants such that the function is continuous.

3. Class-3 Functions : functions with discontinuities in D_2. Out-of-sight path loss models studied in [58], where there is a sudden drop in the power due to blockages in urban/sub-urban areas with buildings provide examples for this class of functions. An example of such a path loss

System Parameters	λ_b Density of base stations p Transmit power	w System bandwidth $l(\mathbf{y})$ Path loss
Spatial Processes	Model	Gaussian Limits
Interference	$\mathcal{J}_{\lambda_b}$; $\mathcal{J}^c_{\lambda_b}$(re-scaled)	$\hat{\mathcal{J}}_{\lambda_b}$; $\hat{\mathcal{J}}^c$ (re-scaled)
Shannon rate	$\mathcal{S}_{\lambda_b}$	$\hat{\mathcal{S}}_{\lambda_b}$
Temporal Processes	Model	Gaussian Limits
Interference	J_{λ_b}	$\hat{J}_{\lambda_b}$;$\hat{J}^c$ (re-scaled)
Shannon rate	S_{λ_b}	$\hat{S}_{\lambda_b}$

Table 3.1: Key Parameters and Processes.

function is:

$$l_3(\mathbf{y}) = \begin{cases} 1 & \text{for } -r_0 \leq y_1 < r_0, -r_0 \leq y_2 < r_0, \\ 0 & \text{for otherwise.} \end{cases} \tag{3.5}$$

A discussion of the approximation of radial discontinuous functions by functions that belong to the D_2 class of functions is given in the Appendix 3.7.1.

3.3 Scaling Limit of the Interference Field

Under our modeling assumptions we have $\mathbb{E}[\mathcal{I}_{\lambda_b}(0)] = \lambda_b p \int_{\mathbb{R}^2} l(\mathbf{y}) d\mathbf{y} < \infty$, so we can consider the following re-scaling of the interference field:

$$\mathcal{I}^c_{\lambda_b}(\mathbf{y}) = \frac{1}{\sqrt{\lambda_b}}(\mathcal{I}_{\lambda_b}(\mathbf{y}) - \mathbb{E}[\mathcal{I}_{\lambda_b}(0)]). \tag{3.6}$$

It is well known that as $\lambda_b \to \infty$, the scaled field $\mathcal{I}^c_{\lambda_b}$ converges to a stationary Gaussian random field $\hat{\mathcal{I}}^c$. Given two locations $\mathbf{z}_1, \mathbf{z}_2 \in \mathbb{R}^2$, the

covariance kernel $c(\mathbf{z}_1, \mathbf{z}_2)$ depends only on the Euclidean distance $t = ||\mathbf{z}_1 - \mathbf{z}_2||$ and is given by:

$$c(t) = \mathbb{E}[\hat{\mathcal{I}}^c(\mathbf{z}_i)\hat{\mathcal{I}}^c(\mathbf{z}_j)] = \int_{\mathbb{R}^2} p^2 l(\mathbf{y} - \mathbf{z}_i) l(\mathbf{y} - \mathbf{z}_j) dy. \tag{3.7}$$

Theorem 13. *For bounded path loss functions that satisfy Lipschitz conditon, consider the re-scaling of the interference field, $\mathcal{I}_{\lambda_b}$ given by:*

$$\mathcal{I}^c_{\lambda_b}(\boldsymbol{y}) = \frac{1}{\sqrt{\lambda_b}}(\mathcal{I}_{\lambda_b}(\boldsymbol{y}) - \mathbb{E}[\mathcal{I}_{\lambda_b}(0)]). \tag{3.8}$$

Then, as $\lambda_b \to \infty$, $\mathcal{I}^c_{\lambda_b}$ converges weakly to a stationary Gaussian random field, $\hat{\mathcal{I}}^c$. Further, in the limit the expectation of the interference field scales with λ_b as

$$\mathbb{E}[\mathcal{I}_{\lambda_b}(0)] = \lambda_b \kappa, \tag{3.9}$$

where

$$\kappa = p \int_{\mathbb{R}^2} l(\boldsymbol{y}) d\boldsymbol{y}. \tag{3.10}$$

In particular, we have convergence in the sense of finite dimensional distribution, i.e., for $\mathbf{z}_1, \mathbf{z}_2, ..., \mathbf{z}_n \in \mathbb{R}^2$, the random vector $(\mathcal{I}^c_{\lambda_b}(\mathbf{z}_1), \mathcal{I}^c_{\lambda_b}(\mathbf{z}_2), ...\mathcal{I}^c_{\lambda_b}(\mathbf{z}_n))$ converges to a centered Gaussian vector $(\hat{\mathcal{I}}^c(\mathbf{z}_1), \hat{\mathcal{I}}^c(\mathbf{z}_2), ...\hat{\mathcal{I}}^c(\mathbf{z}_n))$ for any $n \geq 1$ with covariance kernel, $c(t)$ given by (3.7). In addition, the tightness condition given in [56] can be shown to hold which in turn suffices to show weak convergence of the rescaled interference fields. Proof is given in Appendix 3.7.2.

r_0 **values in meters**	**Base station density(** λ_b**) per Km**^{2}
10	3×10^4
25	10^4
50	8×10^3
100	3×10^2

Table 3.2: The minimum density of base stations such that the scaled interference is a Gaussian field for various values of r_0 of the considered path loss function.

For bounded path loss funtions in the D_2 functional space as defined in [56], we prove the finite dimensional convergence as above, but verify the tightness condition with the help of simulations.

We would like to characterize the interference and the Shannon rate fields. For this we use the above scaling to approximate the interference field. Given the central limit Gaussian field $(\hat{\mathcal{J}}^c(\mathbf{y}), \mathbf{y} \in \mathbb{R}^2)$, for a large values of λ_b, from (3.8), the interference field at location $\mathbf{y}$, $\mathcal{J}_{\lambda_b}(\mathbf{y})$, can be approximated as follows:

$$\mathcal{J}_{\lambda_b}(\mathbf{y}) \sim \sqrt{\lambda_b}\hat{\mathcal{J}}^c(\mathbf{y}) + \mathbb{E}[\mathcal{J}_{\lambda_b}(0)] + o(\sqrt{\lambda_b}). \tag{3.11}$$

Since the expectation of the interference scales linearly with λ_b as seen in Theorem 13, the Gaussian approximation for the interference field $\hat{\mathcal{J}}_{\lambda_b} = (\hat{\mathcal{J}}_{\lambda_b}(\mathbf{y}), \mathbf{y} \in \mathbb{R}^2)$ is given by:

$$\hat{\mathcal{J}}_{\lambda_b}(\mathbf{y}) = \sqrt{\lambda_b}\hat{\mathcal{J}}^c(\mathbf{y}) + \lambda_b\kappa. \tag{3.12}$$

Let us illustrate the convergence of the interference field considering a dual-slope path-loss model of Class-2 (3.4) and for $\lambda_b = 10^4$. The Kolmogorov-

Smirnov test (K-S test) is used to compare the marginal empirical CDF with the Gaussian cummulative distribution. For a given r_0 value in the considered path loss function, there exists a minimum value for the base station density, λ_b, such that the scaled interference is a Gaussian field. Using the K-S test, Table 3.2 lists the minimum density needed for various values of r_0. Fig 3.1 illustrates the comparison of the marginal CDF of the scaled interference field $\mathcal{J}^c_{\lambda_b}(0)$ with the CDF of the Gaussian random variable. We also compare the empirical mean of the interference field $\mathcal{J}_{\lambda_b}(0)$ with the approximate mean given in (3.9) and the approximate mean is within 99% confidence interval.

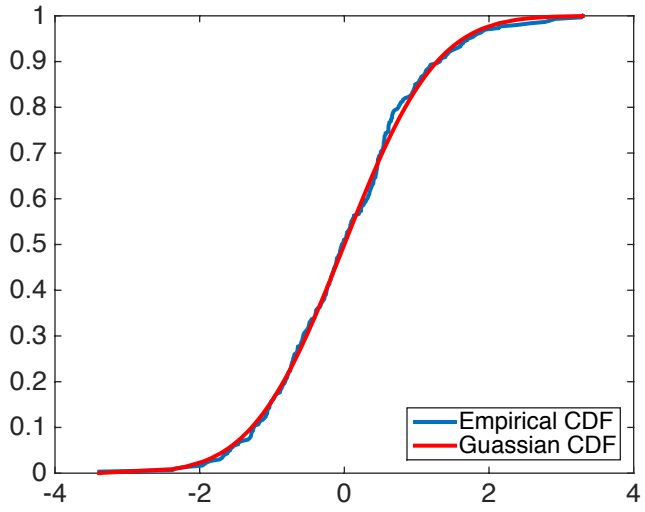


Figure 3.1: Comparing the marginal empirical CDF of $\mathcal{J}^c_{\lambda_b}(0)$ with CDF of a Gaussian.

Now, we focus on certain fundamental questions about the Gaussian field $(\hat{\mathcal{J}}^c(\mathbf{y}), \mathbf{y} \in \mathbb{R}^2)$, such as its continuity and differentiability. Recall that for

a Gaussian field, these are determined by the mean and covariance kernel given in (3.7). We leverage some well known results regarding the continuity and differentiability of Gaussian fields as stated in the lemmas given in Appendix 3.7.3. The following theorem states the result for the path loss functions of Class-1.

Theorem 14. *For Class-1 radial path loss functions, if for some $\eta > 0$ and all $t \in [-\eta, \eta]$, the following integral in polar coordinates*

$$\int_0^{2\pi} \int_0^{\infty} l(r) l'(\sqrt{r^2 + t^2 - 2rt\cos(\theta)}) \times \left(\frac{t - r\cos(\theta)}{\sqrt{r^2 + t^2 - 2rt\cos(\theta)}} \right) r dr d\theta \tag{3.13}$$

is uniformly convergent, and in addition the covariance kernel $c(t)$, defined in (3.7)*, is convergent, then, $c(t)$ is continuous and twice differentiable. Thus, the limiting Gaussian field, $(\hat{\mathcal{J}}^c(\boldsymbol{y}), \boldsymbol{y} \in \mathbb{R}^2)$ is mean square and almost surely continuous and differentiable.*

The proof is given in Appendix 3.7.4.

Note that the covariance kernel $c(t)$ is symmetric about $t = 0$ for all considered path loss functions. Thus the limiting Gaussian fields are always mean square continuous. However when relaxing the conditions on the path loss functions given in the above theorem, mean square differentiability, sample path continuity and differentiability may not hold. We illustrate the math involved in verifying the continuity of the fields for radial path loss functions that belong to Class-2 in Appendix 3.7.5.

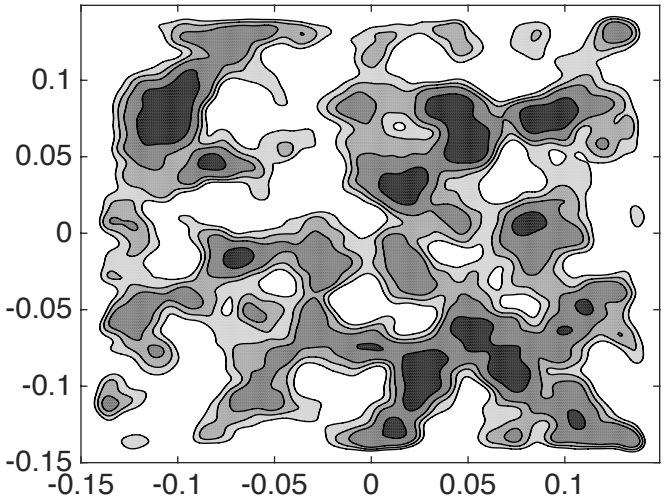


Figure 3.2: Level sets of the limiting Gaussian field, $\hat{\mathcal{J}}^c$, of Class-1 path loss function for different thresholds.

Figure 3.2 and 3.3 illustrate the level sets of the limiting Gaussian field for Class-1 and Class-3 radial path loss functions respectively. One can notice an increase in the spatial variability for the discontinuous path loss functions, i.e., the limiting Gaussian field has high fluctuations with respect to space, which stems from the no-where differentiability property of the field.

3.4 Shannon Rate Field and its Applications

Recall that the Shannon rate field, $(\mathcal{S}_{\lambda_b}(\mathbf{y}), \mathbf{y} \in \mathbb{R}^2)$, is defined as:

$$\mathcal{S}_{\lambda_b}(\mathbf{y}) = w \log\left(1 + \frac{p * l(\mathbf{y} - \mathbf{X}_{\lambda_b}(\mathbf{y}))}{\mathcal{J}_{\lambda_b}(\mathbf{y})}\right). \tag{3.14}$$

By contrast with the traditional polynomial path loss model, for bounded

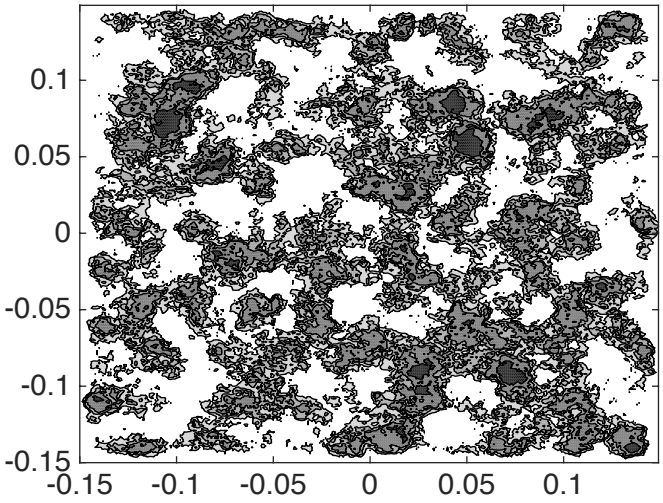


Figure 3.3: Level sets of the limiting Gaussian field, $\hat{\mathcal{J}}^c$, of Class-3 path loss function for different thresholds.

path loss models, the received signal power does not grow to infinity with densification as users get closer to their associated base stations, but the interference keeps growing. Thus, in this regime, the Shannon rate field decreases to zero as the base station density grows. A regime of interest in this setting is one where one densifies and increases the operational bandwidth of the system. In this section, we study this regime obtaining a Gaussian approximation for the Shannon rate field and show its applicability in studying the variability of the spatial average rate and cost of backhauling in dense networks.

3.4.1 Gaussian Approximation for Shannon Rate Field

Under the bounded path loss models given in (3.4), as λ_b tends to ∞, $l(\mathbf{y}-\mathbf{X}_{\lambda_b}(\mathbf{y}))$tends to 1 almost surely, i.e., all users eventually have good signal

path to their closest or more generally the best among the set of possible base stations they can associate with. For simplicity, we then assume the signal path loss is 1. However the interference will be high, thus we can approximate the Shannon rate field as:

$$w \log(1 + p/\mathcal{J}_{\lambda_b}(\mathbf{y})) \sim \frac{wp}{\mathcal{J}_{\lambda_b}(\mathbf{y})}. \tag{3.15}$$

Now based on our characterization of the interference limit as in (3.12) we can approximate the Shannon rate field as:

$$\mathcal{S}_{\lambda_b}(\mathbf{y}) \sim \frac{wp}{\sqrt{\lambda_b}\hat{\mathcal{J}}^c(\mathbf{y}) + \lambda_b \kappa},$$

where κ is defined in (3.10).

Using a Taylor series expansion, one obtains

$$\mathcal{S}_{\lambda_b}(\mathbf{y}) \sim \frac{wp}{\kappa\lambda_b} - \frac{wp\sqrt{\lambda_b}\hat{\mathcal{J}}^c(\mathbf{y})}{\kappa^2\lambda_b^2} + o(1/\lambda_b^{3/2}). \tag{3.16}$$

Thus to compensate for the increase in interference, the Shannon rate must be scaled by a system bandwidth w scaling linearly in λ_b. We let $w = a\lambda_b$ and define our Gaussian model for the Shannon rate field as follows $\hat{\mathcal{S}}_{\lambda_b} = (\hat{\mathcal{S}}_{\lambda_b}(\mathbf{y}), \mathbf{y} \in \mathbf{R}^2)$ where

$$\hat{\mathcal{S}}_{\lambda_b}(\mathbf{y}) = \frac{ap}{\kappa} - \frac{ap}{\kappa^2\sqrt{\lambda_b}}\hat{\mathcal{J}}^c(\mathbf{y}), \tag{3.17}$$

with mean the covariance kernels given by:

$$\begin{aligned} E[\hat{\mathcal{S}}_{\lambda_b}(\mathbf{y})] &= \mu = \frac{ap}{\kappa}, \\ \mathrm{Cov}(\hat{\mathcal{S}}_{\lambda_b}(\mathbf{y}), \hat{\mathcal{S}}_{\lambda_b}(\mathbf{x})) &= \frac{\mu^2}{\lambda_b\kappa^2} c(||\mathbf{y} - \mathbf{x}|)), \end{aligned}$$

where $c(.)$ is given in (3.7).

Note that the variability of the field decreases with the intensity of base stations λ_b. Fig. 3.4 exhibits the marginal empirical CDF of the Shannon rate field along with that of the Gaussian approximation for the case where $\lambda_b = 10^4$ per Km^2 and dual-slope path loss function as in (3.4). As can be seen the match is very good. Similar observations apply in looking at the finite dimensional distributions of these processes.

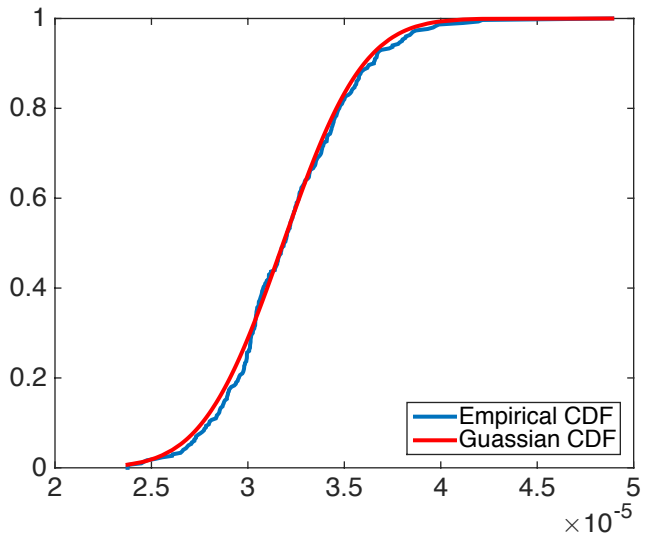


Figure 3.4: Comparing the marginal empirical CDF of the Shannon rate with a Gaussian.

3.4.2 Variability of the Spatial Average Rate (SAR)

Given a Gaussian field such as $\hat{\mathcal{S}}_{\lambda_b} = (\hat{\mathcal{S}}_{\lambda_b}(\mathbf{y}), \mathbf{y} \in \mathbf{R}^2)$, one can now consider various relevant functions of the spatial process. For example, below

we define the spatial average rate over a fixed region.

Definition 6. *The* spatial average *of the Shannon rate field* $\hat{\mathbb{S}}_{\lambda_b}$ *over a set* $A \subset \mathbb{R}^2$ *is defined by*

$$X_{\lambda_b}(A) = \frac{1}{|A|} \int_A \hat{\mathbb{S}}_{\lambda_b}(\boldsymbol{y}) d\boldsymbol{y},$$

where $|A|$ *denotes the area of* A*.*

It follows immediately from the properties of Gaussian processes that $X_{\lambda_b}(A)$ is Gaussian such that

$$\begin{aligned} \mathbb{E}[X_{\lambda_b}(A)] &= \mu, \\ \mathrm{Var}(X_{\lambda_b}(A)) &= \frac{\mu^2}{\kappa^2 \lambda_b |A|^2} \int_A \int_A c(||\mathbf{y} - \mathbf{z}||) d\mathbf{y} d\mathbf{z}, \end{aligned}$$

where, $\kappa = p \int_{\mathbb{R}^2} l(\mathbf{y}) d\mathbf{y}$.

For a fixed region A one might ask how densification will impact variability in the spatial average rate. Our analysis suggests the standard deviation is inversely proportional to $\sqrt{\lambda_b}$, i.e., leads to *concentration* in the rates users will see. Fig 3.5 illustrates such decreases in variability for Class 2 and 3 path loss models. Perhaps as expected, scenarios with more discontinuous path loss characteristics see higher variability but still similar decays.

3.4.3 Backhaul Capacity Dimensioning

The cost and provisioning of backhauling resources is one of the key issues associated with deploying dense networks. In this subsection we shall study how densification impacts the cost of backhauling, leveraging again functionals of our Gaussian model for the Shannon rate field.

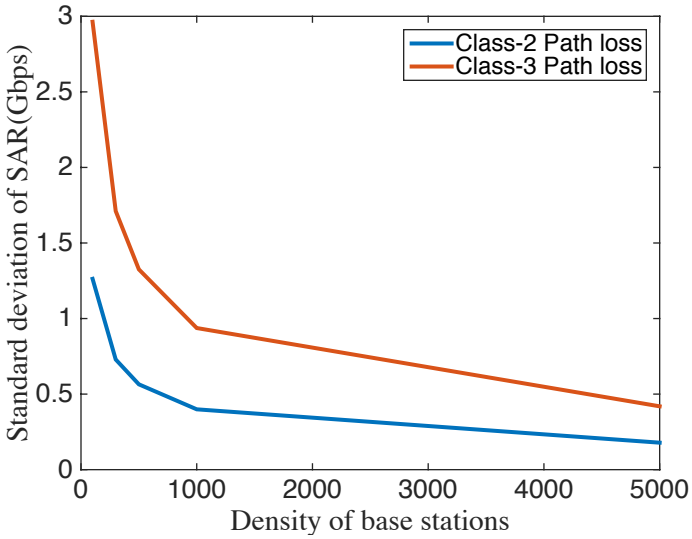


Figure 3.5: Standard deviation of the spatial average rate with increasing density of base stations.

We consider a simple backhauling infrastructure based on a grid tessellation, where each cell (square) is associated with a gateway which provides backhauling for the users in its cell. We assume that the spatial density of users the network serves grows along with the density of base stations. We consider this setting as one of the aims of densification is to provide individual users high throughput which is achieved by cell-splitting gain. Although more general models could be considered, we shall assume that backhauling technology is such that each gateway can handle roughly a fixed number of users say m. Hence, as we density, the cells' area $|A_{\lambda_b}|$ decreases inversely proportional to the base station density i.e., $|A_{\lambda_b}| = \frac{m}{\lambda_b}$. We assume that neither the link from the base station to the gateway nor the backhaul to the Internet is a bot-

tleneck. That is, base stations are connected to the gateway via high capacity links, e.g., mmWave, links. Here the key question is about the capacity that should be provisioned from the gateway to the Internet.

There are various sources of variability which impact the provisioning of the gateway to the Internet backhaul capacity: (1) variability in the peak rate of users, (2) correlations amongst users' peak rates, (3) variability in the number of active users, and (4) sharing of base station resources by one or more users (which limits their peak rate).

Let us first ignore the impact of variability in the number of users. To that end, we shall consider user locations Φ_g corresponding to a grid with density $\lambda_u = \lambda_b$. This scenario might correspond to a deterministic deployment e.g., a video surveillance system with a fixed set of active users. Further, ignoring the sharing of base station resources, we model the aggregate peak rate requirement at a typical gateway cell for a base station density λ_b as: $R_{\lambda_b}(\Phi_g) = \sum_{\mathbf{Y}_j \in \Phi_g \cap A_{\lambda_b}} \hat{\mathcal{S}}_{\lambda_b}(\mathbf{Y}_j)$.

Lemma 2. *The gateway capacity $\rho(\delta)$ capable of serving the aggregate peak rate $R_{\lambda_b}(\Phi_g)$ with overflow probability δ is given by*

$$\rho(\delta) = argmin_{\rho}\{\rho \mid \mathbb{P}(R_{\lambda_b}(\Phi_u) \geq \rho) \leq \delta\}, \tag{3.18}$$

where, $\rho(\delta) = \tilde{\mu} + Q^{-1}(\delta)\tilde{\sigma}$, $Q^{-1}(x) = \sqrt{2}erf^{-1}(1-2x)$, $\tilde{\mu} = E[R_{\lambda_b}(\Phi_g)] = \mu m$, and $\tilde{\sigma}^2 = Var(R_{\lambda_b}(\Phi_g))$ which is given by

$$\frac{\mu^2 m}{\kappa^2 \lambda_b} c(0) + \frac{\mu^2}{\kappa^2 \lambda_b} \sum_{i,j=1, i \neq j}^{m} c(||\boldsymbol{g}_{\lambda_b}^{(i)} - \boldsymbol{g}_{\lambda_b}^{(j)}||),$$

with $\boldsymbol{g}_{\lambda_b}^{(j)}, j = 1, \ldots m$ *corresponding to the* m *grid points in the gateway's square cell.*

Proof. In our grid model, the aggregate peak rate, $R_{\lambda_b}(\Phi_g)$, is the sum of m jointly Gaussian (positively correlated) random variables associated with the grid locations Φ_g in the gateway cell. Thus, $R_{\lambda_b}(\Phi_u)$ is Gaussian $(\tilde{\mu}, \tilde{\sigma}^2)$ and the result in the Lemma follows.

Note that the first term in the variance captures the spatial variations in the users' peak rate while the second term captures correlations amongst the users peak rates. See the Appendix 3.7.6 for a derivation of the above variance formula. □

Remark 3. *(Variability in number of users.) We shall consider the scenario where users' locations* Φ_r *correspond to a Poisson process with intensity* $\lambda_u = \lambda_b$*, which model variations in both the number and locations of active users. Let* $R_{\lambda_b}(\Phi_r)$ *correspond to a random sum of random variables corresponding to a Poisson distributed number of users in the gateway cell, having Gaussian peak rates which are correlated. Such a mixture of Gaussian random variables is no longer Gaussian but is reasonably well approximated. So we propose to approximate* $R_{\lambda_b}(\Phi_r)$ *as a Gaussian Random variable with the same mean as* $R_{\lambda_b}(\Phi_g)$ *and with*

$$Var(R_{\lambda_b}(\Phi_r)) = \mu^2 m + \frac{\mu^2 m}{\kappa^2 \lambda_b} c(0) + \frac{\mu^2 \lambda_b}{\kappa^2} \int_{A_{\lambda_b}} \int_{A_{\lambda_b}} c(||\boldsymbol{y} - \boldsymbol{z}||) d\boldsymbol{y} d\boldsymbol{z}.$$

See the Appendix 3.7.6 for a derivation of the above variance formula. Fig 3.6 illustrates the empirical distribution of the Poisson integral with that of Gaussian random variable to validate this approximation.

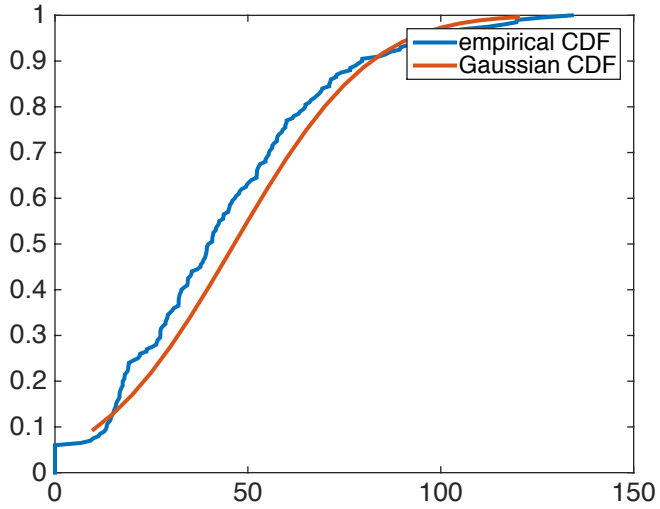


Figure 3.6: Comparing the empirical distribution of the sum rate with a Gaussian distribution with empirical mean and variance.

The first term in the above variance captures variability in the number of users, the second is associated with variability in user's peak rate while the third again captures correlations amongst the users peak rates.

In addition to determining the required backhaul capacity for the above two scenarios, Grid and Random, we can also determine the capacity one would provision if one ignored the terms in their variance corresponding to positive spatial correlations in users peak rates. We refer to the latter as Grid-simple and Random-simple.

λ_b		**Class-1**	**Class-2**	**Class-3**
100	Φ_g	8.9%	22%	19.2%
	Φ_r	0.8%	8.1%	9.6%
300	Φ_g	5.3%	21.2%	22.9%
	Φ_r	0.2%	5.6%	8.5%
500	Φ_g	4.1%	19.2%	22.7%
	Φ_r	0.1%	4.3%	7.3%

Table 3.3: The values of the relative increase in the required backhaul capacity for various classes of path loss functions.

We evaluated the required backhaul capacity according to Lemma 1, for $\delta = 0.01, \lambda_b|A_{\lambda_b}| = m = 20$ and dual slope path loss with $r_0 = 100$m for all the above mentioned scenarios and Fig 3.7 exhibits the comparison of the required backhaul capacity with increasing density, λ_b. We have the following observations:

- Since a single gateway serves approximately a fixed number of base stations, m, the capacity can be viewed as the required backhaul capacity per unit base station. Thus, the cost of providing backhaul capacity decreases with densification due to decrease in the variance of the Shannon rate.

- As seen in Fig.3.7, the positive rate correlations impact the capacity requirements differently for random and grid users. The relative increase in the capacity ranges from 22% - 19.2% for grid users and 8.1% - 4.3% for random users. The user variability dominates the variability due to correlations. Thus the relative increase in the required capacity is higher

for grid users.

- The positive rate correlations also impact differently for various path loss models and Table 3.3 states the values of the relative increase in the required backhaul capacity for various classes of path loss functions.

- The required capacity is higher for random users due to the additional contribution of user variability.

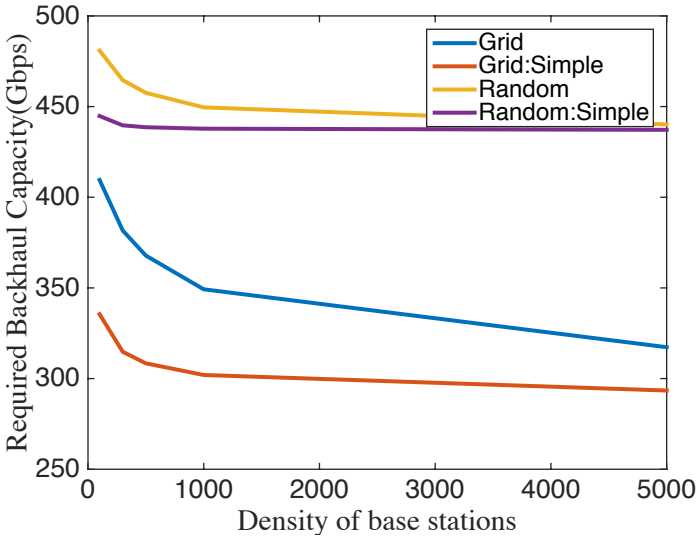


Figure 3.7: Required backhaul capacity with increasing density of base stations.

3.5 Temporal Characteristics of the Shannon Rate Process

Definition 7. *Given a stationary spatial field, such as the interference field* $\mathcal{J}_{\lambda_b}$*, the temporal stochastic process seen by a mobile moving at a constant*

velocity, v along a straight line, $(J_{\lambda_b}(t), t > 0)$ is defined as:

$$J_{\lambda_b}(t) = \mathcal{J}_{\lambda_b}(\boldsymbol{y}_v(t)), \tag{3.19}$$

where $\boldsymbol{y}_v(t) = (vt, 0)$.

Similarly, we can define the temporal Shannon rate process, $(S_{\lambda_b}(t), t > 0)$. Given the fields are stationary and isotropic, without loss of generality, we can assume that the user is moving along the x-axis starting from the origin at time $t = 0$. In a first step, consider that the mobile user is moving at a fixed unit velocity.

Given the asymptotic characterization of the interference and Shannon rate fields as Gaussian fields, one can asymptotically characterize the above stochastic processes as stationary Gaussian processes, $(\hat{J}_{\lambda_b}(t), t > 0)$ and $(\hat{S}_{\lambda_b}(t), t > 0)$. The continuity and differentiability properties of the processes follow immediately from those of the fields.

Note that, as the mobile moves through space, the variations in the rate it experiences depend on the path loss functions. For Class-1 functions, we have smooth variations in the interference and rate i.e., differentiable processes. For Class-3 functions we have no-where differentiable processes with high variations as illustrated in Fig 3.8. To analyze the high temporal variations in the rate, we characterize them with the help of Hölder exponents [12].

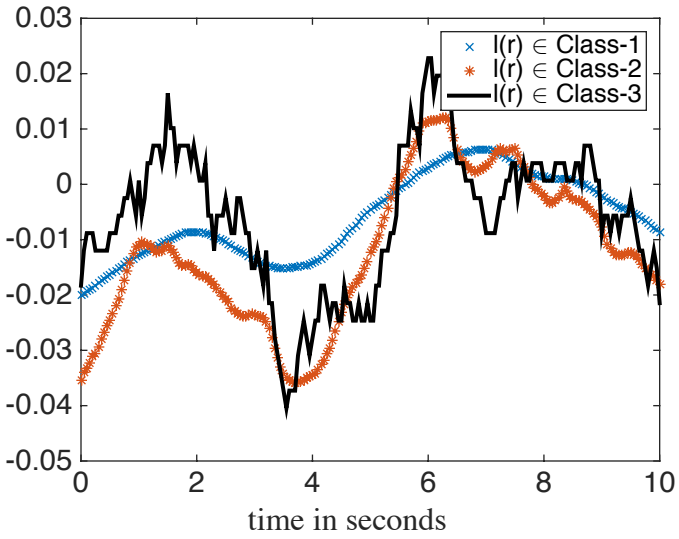


Figure 3.8: Sample path of Interference processes for various path loss functions.

3.5.1 Hölder Exponents

Definition 8. *[12] A function $g : \mathbb{R} \to \mathbb{R}$ has a Hölder exponent α if there exists a constant k such that for any $\hat{k} > k$ and all sufficiently small h,*

$$|g(t+h) - g(t)| < \hat{k}|h|^{\alpha},$$

uniformly with respect to all t lying in any finite interval. We then say that $g(t) \in H(\alpha, k)$.

The Hölder exponent, $0 < \alpha < 1$, provides a measure of local path-irregularity or roughness: sample paths exhibit more and more variability as α decreases from 1 to 0.

Lemma 3. *[12] For a given stationary continuous Gaussian process* $(\hat{J}^c(t), t \geq 0)$ *with covariance kernel,* $c(t)$ *if, for all sufficiently small* h*,*

$$\mathbb{E}[(\hat{J}^c(t+h) - \hat{J}^c(t))^2] = 2(c(0) - c(h)) \leq \frac{k|h|^{2\alpha}}{|\log|h||}, \tag{3.20}$$

then almost surely all sample paths of $\hat{J}^c(t)$ *belong to* $H(\alpha, \sqrt{2k}/k_\alpha)$*, where* $k_\alpha = \frac{2^\alpha - 1}{2^{1+2\alpha}}$*. Further, the Hölder exponent of the stochastic process as defined in [12] is* α*.*

Remark 4. *Consider the path loss model defined in* (3.47) *with* $f(r_0) = 0$*. The interference experienced by the mobile user moving with unit velocity along a straight line is a stochastic process that is equivalent to the number of users in a* $M/GI/\infty$ *queue, with service rate given by*

$$f_{W^{(\gamma)}}(s) = \begin{cases} \frac{s}{2r_\gamma\sqrt{4r_\gamma^2 - s^2}} & \text{for } s \in [0, 2r_\gamma], \\ 0 & \text{otherwise} \end{cases} \tag{3.21}$$

(see [59]). The limiting process of the re-scaled version of the process tracking the number of users in an $M/GI/\infty$ *queue, is a Gaussian process that is almost sure continuous, no where differentiable with covariance kernel given by* $c'(t) = \int_0^\infty [1 - F_{W^\gamma}(t-u)]du$ *([60], [61]), which is the same as the covariance given by* (3.7) *for the given path loss function. Thus, the Hölder exponent of the process is 1/2.*

For certain Class-2 and Class-3 path loss functions of interest, we could verify that they satisfy the above condition by numerically evaluating the covariance kernel.

3.5.2 Time Scales in Adaptive Modulation and Coding

In our environment, i.e., dense networks with the bounded path loss functions considered here and no fading, the signal power is asymptotically a constant, and the variability in the Shannon rate is primarily due to variations in the interference. In this section, we discuss ways to cope with such interference variations in the context of adaptive modulation and coding.

Adaptive modulation and coding is a technique used to adapt to variations in signal quality (SINR), where it dynamically selects the best modulation and coding scheme (MCS) based on estimates of current conditions. For simplicity, in our model, we assume uniform binning of the Shannon rate itself with bin size Δ. Let $b_0, b_1, \ldots, b_n$ be the discrete rates defining the bins, where bin i is associated with rate b_i and $|b_i - b_{i-1}| = \Delta$. This is a simplified model, as in practice, the range of the estimated channel state information (CSI) (e.g., SINR) is divided into non-uniform bins, each corresponding to a MCS, which are then mapped to a transmission rate by a non-linear function.

We now explore the timescales on which the adaptive modulation and coding (AMC) should operate. Slower rate of adaptation leads to difficulties in keeping up with the local variations. The user may experience conditions worse than those required for the selected rate, which should be avoided. At the same time a significant amount of overhead is involved in estimating the interference power and selecting a new rate. Thus one should try to limit the rate of adaptations.

We use our model to provide an understanding of the rate of adaptation and its dependence on various system parameters. Assume that adaptive coding takes place periodically every h seconds. Namely, every h seconds, the transmitter selects a particular rate, $b_\sigma \in \{b_1, \ldots, b_n\}$ bits/sec based on the estimate of the instantaneous Shannon rate σ at the given time. The selected rate b_σ is then used for the next h seconds.

In order to cope with variations, we consider a conservative approach where, if the current rate σ belongs to bin i, then we pick a code rate corresponding to bin $(i-1)^+$, i.e., $b_\sigma = b_{(i-1)^+}$. Then we have the following theorem which gives a way to choose h such that the chance that the selected modulation and coding rate is fine for the next h seconds.

Theorem 15. *For our adaptive modulation and coding model, if the centered Gaussian process, $(\hat{J}^c(t), t \geq 0)$ satisfies the condition given in Lemma 3, then for all $|h| \leq g_{\lambda_b}(\alpha, k)$ with:*

$$g_{\lambda_b}(\alpha, k) = \left(\frac{\Delta \kappa^2 \sqrt{\lambda_b}}{ap\hat{k}} \right)^{1/\alpha}, \tag{3.22}$$

and $\hat{k}$ any number such that $\hat{k} > \frac{\sqrt{2k}}{k_\alpha}$, where $k_\alpha = \frac{2^\alpha - 1}{2^{1+2\alpha}}$, almost surely, all the sample paths satisfy $|\hat{S}_{\lambda_b}(t+h) - \hat{S}_{\lambda_b}(t)| < \Delta$, where $\hat{S}_{\lambda_b}$ is the Gaussian approximation of the Shannon rate process.

Proof. For h small, from (3.17), $|\hat{S}_{\lambda_b}(t+h) - \hat{S}_{\lambda_b}(t)| < \Delta$ if and only if $|\hat{J}^c(t+h) - \hat{J}^c(t)| < \frac{\Delta \kappa^2 \sqrt{\lambda_b}}{ap}$. If the limiting Gaussian interference process, $\hat{J}^c(t)$, satisfies the conditions of Lemma 3, then for all sufficiently small h and $\hat{k} >$

$\sqrt{2k}/k_\alpha$ almost surely, all sample paths satisfy $|\hat{J}^c(t+h) - \hat{J}^c(t))| < \hat{k}|h|^\alpha$. Thus, for all $|h| \leq g_{\lambda_b}(\alpha, k)$, almost surely for all the sample paths, $|\hat{S}_{\lambda_b}(t+h) - \hat{S}_{\lambda_b}(t)| < \Delta$. □

Role of velocity. Let us now assume that the mobile user is moving with constant velocity v instead of unit velocity. From Definition 7, we have

$$|\hat{J}^c(t+h) - \hat{J}^c(t))| = |\hat{\mathcal{J}}^c(\mathbf{y}_v(t+h)) - \hat{\mathcal{J}}^c(\mathbf{y}_v(t))|, \tag{3.23}$$

where $\mathbf{y}_v(t) = (vt, 0)$. Further, since bin size has to be in the same order of magnitude as the Shannon rate, it makes sense to assume that the bin size is a fraction of the mean Shannon rate, i.e., $\Delta = \frac{1}{\eta}\mathbb{E}[S_{\lambda_b}(0)]$ for some $\eta \in \mathbb{R}^+$ e.g., $\eta = 10$. Then, the function $g_{\lambda_b}(\alpha, k, v)$ is now given by:

$$g_{\lambda_b}(\alpha, k) = \frac{1}{v}\left(\frac{\kappa\sqrt{\lambda_b}}{\eta\hat{k}}\right)^{1/\alpha}. \tag{3.24}$$

The time period at which AMC should operate at primarily depends on: (1) the intensity of base stations (λ_b), (2) the velocity (v), and (3) the path loss models through $\kappa, \hat{k}, \alpha$. We study these various dependencies by considering a specific set of values for parameters: Intensity of base stations, $\lambda_b = 300, 1000$ per Km^2; Bandwidth, $w = 900$ MHz; transmitted power, $p = 1$ Watt and velocity, $v =$ from 1 to 10 m/s.

Further, we consider the dual slope path loss function of Class-2 and a discontinuous path loss function. We then numerically estimate the Hölder exponents to be 1 and 0.5 with constant values k of 5 and 50 respectively. Then, for $\hat{k} = \frac{\sqrt{2k}}{k_\alpha}$, the function $g_{\lambda_b}(\alpha, k)$ is given in Table 3.4.

λ_b	v m/s	**Class-2**	**Class-3**
300	1	$\sim$ ms	$\sim 10\mu$ s
	10	$\sim 10^{-1}$ ms	$\sim \mu$s
1000	1	$\sim 10^{-1}$ s	$\sim$ ms
	10	$\sim 10^{-2}$ s	$\sim 10^{-1}$ ms

Table 3.4: Time scales determined by Theorem 3, equation (3.24).

From the numerical evaluation and (3.24), we have the following observations:

- We get a lower rate of adaptation when increasing the density of base stations, λ_b, since the variance of the process decreases as studied in the previous section. The magnitude of the rate is due to the fact that function g in (3.24) is polynomial in λ_b which is in per Km^2.
- Increasing the velocity increases the rate of adaptation. For the same set of parameters, if one considers unit velocity we have that $g_{\lambda_b}(\alpha, k) \sim$ ms for Class-2 path loss models.
- Higher variability in the Shannon rate, i.e., lower values of α in case of discontinuous path loss models, leads to higher rates of adaptation.

Table 3.4 illustrates the time scales determined by Theorem 3, equation (3.24), at which adaptive modulation and coding should operate for the above mentioned network parameters. We validated this result with the help of simulations, by considering uniform binning of a fixed bin size Δ of the Shannon rate seen by a user moving at a constant velocity $v = 1$ m/s with $\lambda_b = 300$ and

Time Scale	10μs	1 ms	10^{-1}s
Class-2	0.07	0.09	0.21
Class-3	0.12	0.26	0.5

Table 3.5: Fraction of time periods in error for class-2 and class-3 path loss functions at various time scales.

the network parameters as above. We considered three different time scales, 10μs, 1ms and 10^{-1}s, for adaptive modulation and evaluated the fraction of the time periods that are in error when applying the technique defined above. The simulation values are in agreement with our numerical result given by the Hölder exponent analysis, since the fraction of error is considerably lower for Class-2 path loss functions at 1ms and for Class-3 path loss functions at 10μs as can be seen in Table 3.5. The fraction of error is higher if one considers time scales greater than the values determined by (3.24).

Thus, this provides an understanding of how h scales with different system parameters like the environment through the path loss $l(r)$, density of base stations, λ_b and velocity v.

3.5.3 Level-crossings of the Scaled Interference Process

In this subsection, we study the expected number of up-crossings in an interval of the stationary differentiable Gaussian process $(\hat{J}^c(t), t \geq 0)$. We assume that this process is a.s differentiable. This also gives the expected up-crossings for the approximated interference, $(\hat{J}_{\lambda_b}(t), t \geq 0)$ and Shannon rate processes, $(\hat{S}_{\lambda_b}(t), t \geq 0)$.

Given a threshold u, define the number of up-crossings in an interval $[0,T]$ as $N_u^+[0,T] = \#\{t \in [0,T] : \hat{J}^c(t) = u, \hat{J}^{c'}(t) > 0\}$. Then, the Rice formula for the expected number of upcrossings in the interval is given by ([62])

$$\mathbb{E}[N_u^+[0,T]] = \frac{\sqrt{\omega_2}Te^{-u^2/(2c(0))}}{2\pi\sqrt{c(0)}}, \tag{3.25}$$

where ω_2 is the second spectral moment. Since the Gaussian process is mean square differentiable, the second spectral moment is given by $\omega_2 = -c^{''}(0)$.

Using simulations, we estimated the expected number of up-crossings in an interval for various sample paths of the scaled interference process, $J^c_{\lambda_b}(t)$. The aim is to compare this with the above result for the limiting Gaussian process. Thus, we evaluated Rice formula by calculating the second spectral moment numerically. The simulated mean value is within a 5 percent error margin from the numerically value. Thus, one can expect to use the results for the limiting Gaussian process to work for the original processes. We can also characterize the coverage and outage times i.e., the level crossings of the approximated Gaussian Shannon rate process, $(\hat{S}_{\lambda_b}(t), t \geq 0)$ and its asymptotics using the existing results as in [62], [63].

3.6 Conclusion

By properly rescaling the interference and Shannon rate fields we have characterized their corresponding limiting Gaussian fields in ultra dense settings. This opens the opportunity to apply the rich set of tools and results for Gaussian fields to study dense wireless networks. Their characteristics depend

primarily on the path loss. By taking functions of these fields, one can also shed light on fundamental engineering questions in ultra dense networks such as (1) the role of spatial correlations on backhaul dimensioning and (2) the characteristics of temporal variations mobile users would see and their impact on adaptive modulation and coding. Overall this provides a new approach for the assessment of the fundamental characteristics of densification.

3.7 Appendix

3.7.1 Approximation for Radial D_2 Function

Since functions in D_2 are in a sense "continuous from above with limits from below" [56], discontinuous radial functions do not belong to D_2. Thus, we consider a set K as given in Fig. 3.9, where we approximate the curvature in second and third quadrants with small squares of side length, ϵ. We can then approximate the following radial function:

$$l_3(r) = \begin{cases} 1 & \text{for } r \leq r_0, \\ 0 & \text{for } r > r_0, \end{cases} \tag{3.26}$$

with a D_2 function $l_3(\mathbf{z}) = \chi_K(\mathbf{z})$, where χ denotes the indicator function and K is set as given in Fig. 3.9. Note that lower values of the side length of the squares, ϵ in the set K leads to better approximation of the curvature.

3.7.2 Proof of Theorem 13

Proof. We will use the previous result that the re-scaled shot noise field, in the limit, is a Gaussian field. Let us consider the result as an analogue of

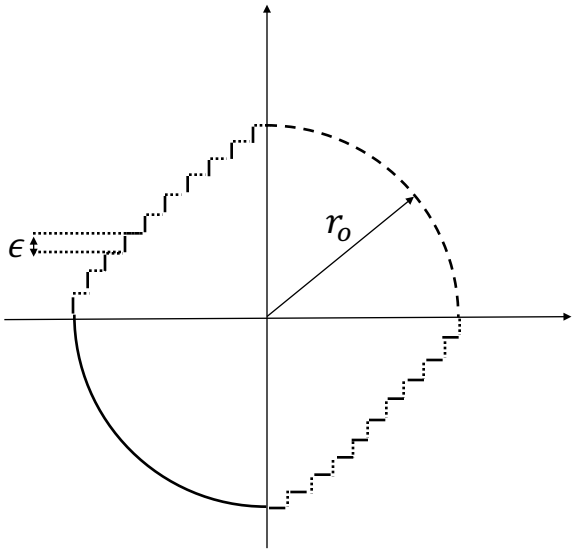


Figure 3.9: The set $K \in \mathbb{R}^2$.

the central limit theorem. Assume that the intensity of the base stations, λ_b is an integer. Then, let $\Phi_i = \{\mathbf{X}_1^i, \mathbf{X}_2^i, ...\}$ for $i = 1, 2, ..\lambda_b$ be Poisson point processes of base stations of intensity 1 and $Y_1(0), Y_2(0), ...Y_{\lambda_b}(0)$ be i.i.d random variables denoting their shot noise field at origin respectively i.e.,

$$Y_i(0) = \sum_{\mathbf{X}_j^i \in \Phi_i} pl(\mathbf{X}_j^i). \tag{3.27}$$

Notice that the shot noise field $\mathcal{I}_{\lambda_b}(0)$ is the sum of the i.i.d random variables $Y_i(0)$ i.e., $\mathcal{I}_{\lambda_b}(0) = \sum_{i=1}^{\lambda_b} Y_i(0)$. Then, from the central limit theorem we have:

$$\lim_{\lambda_b \to \infty} \frac{1}{\sqrt{\lambda_b}} \left(\sum_{i=1}^{\lambda_b} Y_i(0) - \lambda_b \mathbb{E}[Y_i(0)] \right) \to^d \mathcal{N}(0, \sigma^2), \tag{3.28}$$

where $\sigma^2 = \int_{\mathbb{R}^2} p^2 l^2(\mathbf{y}) d\mathbf{y}$ since the LHS in the above equation is the scaled

shot noise field given in (3.6). Let $M(0) = \max_i(\max_j(pl(X_j^i))$. Then, the interference field can also be represented as follows:

$$\mathcal{J}_{\lambda_b}(0) = \sum_{i=1}^{\lambda_b} Y_i(0) - M(0), \tag{3.29}$$

and the scaled interference field can be represented as:

$$\mathcal{J}^c_{\lambda_b}(0) = \frac{1}{\sqrt{\lambda_b}}\bigg(\sum_{i=1}^{\lambda_b} Y_i(0) - \lambda_b\mathbb{E}[Y_i(0)] + (\mathbb{E}[M(0)] - M(0))\bigg). \tag{3.30}$$

Given the bounded path loss model, $M(0)$ is bounded above by a constant p. Thus,

$$-p \leq \mathbb{E}[M(0)] - M(0) \leq p. \tag{3.31}$$

Thus, the scaled interference field can be bounded above and below by:

$$\begin{aligned}\frac{1}{\sqrt{\lambda_b}}(\mathcal{I}_{\lambda_b}(0) - \mathbb{E}[\mathcal{I}_{\lambda_b}(0)] - p) &\leq \frac{1}{\sqrt{\lambda_b}}(\mathcal{J}_{\lambda_b}(0) - \mathbb{E}[\mathcal{J}_{\lambda_b}(0)]) \\ &\leq \frac{1}{\sqrt{\lambda_b}}(\mathcal{I}_{\lambda_b}(0) - \mathbb{E}[\mathcal{I}_{\lambda_b}(0)] + p).\end{aligned} \tag{3.32}$$

Now, as $\lambda_b \to \infty$, both the higher and lower limit converge to Gaussian random variable from (3.28). Thus, $\mathcal{J}^c_{\lambda_b}(0)$ also converges in distribution to Gaussian random variable.

The random vector $Z^{\mathcal{J}} = (\mathcal{I}^c_{\lambda_b}(\mathbf{z}_1), \mathcal{I}^c_{\lambda_b}(\mathbf{z}_2), ..., \mathcal{I}^c_{\lambda_b}(\mathbf{z}_n))$ converges to a Gaussian random vector $(\hat{\mathcal{I}}^c(\mathbf{z}_1), \hat{\mathcal{I}}^c(\mathbf{z}_2), ..., \hat{\mathcal{I}}^c(\mathbf{z}_n))$ with covariance matrix C if for any vector $\mathbf{u} = [u_1, u_2, ...u_n] \in \mathbb{R}^n, n \in \mathbb{N}$, $\mathbf{u}^T Z^{\mathcal{J}} \sim \mathcal{N}(0, \mathbf{u}^T C\mathbf{u})$, where $C_{ij} = c(||\mathbf{z}_i - \mathbf{z}_j||)$ which is given in (3.7). Let us consider the i.i.d random

variables $Z_1^{\mathcal{J}}, Z_2^{\mathcal{J}}, Z_3^{\mathcal{J}}, ...$ defined as follows:

$$Z_i^{\mathcal{J}} = \sum_{j=1}^{n} u_j Y_i(\mathbf{z}_j). \tag{3.33}$$

Notice that the random variable $\mathbf{u}^T Z^{\mathcal{J}}$ is the sum of the i.i.d random variables $Z_i^{\mathcal{J}}$. Then, from multidimensional central limit theorem we have:

$$\lim_{\lambda_b \to \infty} \frac{1}{\sqrt{\lambda_b}} \Bigg(\sum_{i=1}^{\lambda_b} Z_i^{\mathcal{J}} - \lambda_b \mathbb{E}[Z_i^{\mathcal{J}}] \Bigg) \to^d \mathcal{N}(0, \mathbf{u}^T C \mathbf{u}). \tag{3.34}$$

Given the random variables $M(\mathbf{z}_j), j = [1, 2, ..n]$, we have

$$\mathbf{u}^T Z^{\mathcal{J}} = \frac{1}{\sqrt{\lambda_b}} \Bigg(\sum_{i=1}^{\lambda_b} Z_i^{\mathcal{J}} - \lambda_b \mathbb{E}[Z_i^{\mathcal{J}}] + \sum_{j=1}^{n} u_j (\mathbb{E}[M(\mathbf{z}_j)] - M(\mathbf{z}_j)) \Bigg). \tag{3.35}$$

The random variables $M(\mathbf{z}_j), j = [1, 2.., n]$ follow the same bounds as given above, we can bound $\mathbf{u}^T Z^{\mathcal{J}}$ as follows

$$\frac{1}{\sqrt{\lambda_b}} \Bigg(\sum_{i=1}^{\lambda_b} Z_i^{\mathcal{J}} - \lambda_b \mathbb{E}[Z_i^{\mathcal{J}}] - p| \sum_{j=1}^{n} u_j| \Bigg) \leq \mathbf{u}^T Z^{\mathcal{J}} \leq \frac{1}{\sqrt{\lambda_b}} \Bigg(\sum_{i=1}^{\lambda_b} Z_i^{\mathcal{J}} - \lambda_b \mathbb{E}[Z_i^{\mathcal{J}}] + p| \sum_{j=1}^{n} u_j| \Bigg). \tag{3.36}$$

Now, as $\lambda_b \to \infty$, both the higher and lower limit converge to Gaussian random variable from (3.34). Thus, $\mathbf{u}^T Z^{\mathcal{J}}$ also converges in distribution to Gaussian random variable.

Then, $(\mathcal{J}^c_{\lambda_b}(\mathbf{z}_1), \mathcal{J}^c_{\lambda_b}(\mathbf{z}_2), ..., \mathcal{J}^c_{\lambda_b}(\mathbf{z}_n))$ converges to a Gaussian random vector $(\hat{\mathcal{J}}^c(\mathbf{z}_1), \hat{\mathcal{J}}^c(\mathbf{z}_2), ..., \hat{\mathcal{J}}^c(\mathbf{z}_n))$ with covariance kernel c, given by (3.7).

From (3.29), we have that $\mathbb{E}[\mathcal{J}_{\lambda_b}(0)] = \mathbb{E}[\mathcal{I}_{\lambda_b}(0)] - \mathbb{E}[M(0)]$. Recall that the $\mathbb{E}[\mathcal{I}_{\lambda_b}(0)] = \lambda_b p \int_{\mathbb{R}^2} l(\mathbf{y}) d\mathbf{y}$ and that $M(0)$ is bounded above by constant p.

Thus, the result follows from the following bound:

$$\lambda_b p \int_{\mathbb{R}^2} l(\mathbf{y}) d\mathbf{y} - p \leq \mathbb{E}[\mathcal{J}_{\lambda_b}(0)] \leq \lambda_b p \int_{\mathbb{R}^2} l(\mathbf{y}) d\mathbf{y}.$$

In order to show weak convergence of the rescaled fields to the Gaussian field, one must also show the tightness of the limiting sequence. An appropriate tightness condition is given in the Theorem 8 [43]. It requires that the rescaled interference processes be such that there exists a positive measure μ on $[0,1] \times [0,1]$ and λ_0 such that

$$\mathbb{E}\left[(\hat{\mathcal{J}}^c_{\lambda_b}(B))^2 (\hat{\mathcal{J}}^c_{\lambda_b}(C))^2\right] \leq \mu^2(B)\mu^2(C), \tag{3.37}$$

for all pairs of neighboring blocks B and C in $[0,1] \times [0,1]$ and for all $\lambda_b \geq \lambda_0$, where for each block B with ordered corners $\mathbf{x}_1^B, \mathbf{x}_2^B, \mathbf{x}_3^B, \mathbf{x}_4^B$, we have that

$$\hat{\mathcal{J}}^c_{\lambda_b}(B) = \mathcal{J}^c_{\lambda_b}(\mathbf{x}_1^B) - \mathcal{J}^c_{\lambda_b}(\mathbf{x}_2^B) - \mathcal{J}^c_{\lambda_b}(\mathbf{x}_3^B) + \mathcal{J}^c_{\lambda_b}(\mathbf{x}_4^B) \tag{3.38}$$

represents the "increment" of $\mathcal{J}^c_{\lambda_b}$ around B. From [43], we know that the centered shot noise field $\mathcal{I}^c_{\lambda_b}$ satisfies the condition in (3.37). Let $\mathcal{M}_{\lambda_b}$ denote the maximal shot noise field, i.e.,

$$\mathcal{M}_{\lambda_b}(\mathbf{t}) = \sup_{\mathbf{X}_i \in \Phi} l(\mathbf{X}_i - \mathbf{t}). \tag{3.39}$$

Using the Cauchy-Schwarz inequality and the relation $\mathcal{J}^c_{\lambda_b}(\mathbf{t}) = \mathcal{I}^c_{\lambda_b}(\mathbf{t}) - \mathcal{M}^c_{\lambda_b}(\mathbf{t})$, one can show that the protected-shot noise field satisfies the tightness condition if the maximal shot noise field satisfies the condition in (3.37).

Given the characterization of the joint distribution for the maximal shot noise process in [55], we can evalute the expectation of the increment of this field as follows:

$$\mathbb{E}\left[\hat{\mathcal{M}}^{c}_{\lambda_b}(B)\right] = \mathbb{E}\left[l(\mathbf{x}_1^B - \mathbf{R}_1) + l(\mathbf{x}_3^B - \mathbf{R}_3) - l(\mathbf{x}_2^B - \mathbf{R}_2) - l(\mathbf{x}_4^B - \mathbf{R}_4)\right], \tag{3.40}$$

where $\mathbf{R}_i$ is a random variable denoting the location of the closest base station in Φ from location $\mathbf{x}_i^B$. The random variables are not independent since the location of the closest base station for one corner limits the distance to the closest base station for remaining corners. Further, based on the geometry of the associated quantities one can show that they almost surely satisfy the following inequality:

$$|(\mathbf{x}_1^B - \mathbf{R}_1) - (\mathbf{x}_2^B - \mathbf{R}_2)| \leq^{a.s} |\mathbf{x}_1^B - \mathbf{x}_2^B|. \tag{3.41}$$

Using the above inequality, we can bound the increament of maximal shot noise field over a block B by a positive measure of B for path loss functions statisfying the Lipschitz conditon:

$$|l(\mathbf{x}_1^B) - l(\mathbf{x}_2^B)| \leq c|\mathbf{x}_1^B - \mathbf{x}_2^B|, \tag{3.42}$$

for some positive constant c. Hence, we can prove the tightness condition for the protected shot noise fields.

□

3.7.3 Continuity of Gaussian Fields

Here we state the conditions for a generic stationary Gaussian field $(\mathcal{X}(\mathbf{y}), \mathbf{y} \in \mathbb{R}^2)$ with covariance kernel $r(t)$, to be m.s and a.s continuous.

Lemma 4. *[63] The limiting stationary Gaussian field $(\mathcal{X}(\boldsymbol{y}), \boldsymbol{y} \in \mathbb{R}^2)$ is continuous in the mean square sense if and only if its covariance kernel $r(t)$ is continuous at $t = 0$ and is mean square differentiable if and only if its covariance kernel $c(t)$ is twice continuously differentiable in a neighborhood of $t = 0$.*

Lemma 5. *[64] The limiting stationary Gaussian field $(\mathcal{X}(\boldsymbol{y}), \boldsymbol{y} \in \mathbb{R}^2)$ has continuous sample paths if its covariance satisfies*

$$r(0) - r(t) \leq \frac{\xi}{|\log(t)|^{1+\alpha_2}}, \tag{3.43}$$

for some $\alpha_2 > 0$ and some ξ.

3.7.4 Proof of Theorem 14

Proof. We use polar coordinates for our analysis, thus the covariance function given in (3.7) is given by:

$$c(t) = p^2 \int_0^{2\pi} \int_0^{\infty} l(r) l(\sqrt{r^2 + t^2 - 2rt\cos(\theta)}) r dr d\theta. \tag{3.44}$$

Let us define the function $f(r, \theta, t) = l(r) l(\sqrt{r^2 + t^2 - 2rt\cos(\theta)}) r$. The functions $f(r, \theta, t)$ and $\frac{df}{dt}$ are continuous in the domain $0 < r < R, 0 < \theta < 2\pi, -\eta < t < \eta$ for every $R > 0$, since the path loss function is continuous and everywhere differentiable.

If the given two conditions stated are satisfied, we can interchange the differential and the integration. Thus, we have

$$\begin{aligned}\frac{d}{dt}c(t) &= p^2\frac{d}{dt}\int_0^{2\pi}\int_0^{\infty} l(r)l(\sqrt{r^2+t^2-2rt\cos(\theta)})rdrd\theta \\ &= p^2\int_0^{2\pi}\int_0^{\infty} \frac{d}{dt}l(r)l(\sqrt{r^2+t^2-2rt\cos(\theta)})rdrd\theta \\ &= p^2\int_0^{2\pi}\int_0^{\infty} l(r)l^{'}(\sqrt{r^2+t^2-2rt\cos(\theta)})\frac{t-r\cos(\theta)}{\sqrt{r^2+t^2-2rt\cos(\theta)}}rdrd\theta.\end{aligned} \tag{3.45}$$

Thus, $\frac{d}{dt}c(t)$ at $t=0$ is $\int_0^{2\pi}\int_0^{\infty} l(r)l^{'}(r)r\cos(\theta)drd\theta = 0$, since $\int_0^{2\pi}\cos(\theta)d\theta = 0$.

Hence the covariance kernel is twice differentiable. Therefore, $\hat{\mathcal{J}}^c$ is mean square continuous and differentiable with $\hat{\mathcal{J}}^{c^{'}}$ a stationary Gaussian centered field with covariance kernel $-c^{''}(t)$. Moreover,

$$\mathbb{E}\big(\hat{\mathcal{J}}^{c^{'}}(t)-\hat{\mathcal{J}}^{c^{'}}(t^{'})\big)^2\big) = 2(c^{''}(0)-c^{''}(t-t^{'})) \leq 2||l^{'}||_{\infty}||l^{''}||_1|t-t^{'}|, \tag{3.46}$$

thus, $\hat{\mathcal{J}}^{c^{'}}$ and $\hat{\mathcal{J}}^{c}$ are almost surely continuously differentiable(Theorem 1.4.2) [64]. Thus, the Gaussian field is both mean square and sample path continuously differentiable for path loss functions of Class-1. □

3.7.5 Verification of differentiability for a family of path loss function

Let us consider a class of radial path loss functions that are often used in the literature:

$$l(\mathbf{y}) = \begin{cases} 1 & \text{for } ||\mathbf{y}|| \leq r_0, \\ f(||\mathbf{y}||) & \text{for } ||\mathbf{y}|| > r_0. \end{cases} \tag{3.47}$$

We prove that in the case where the path loss function is continuous, i.e., $f(r_0) = 1$, the limiting Gaussian field is mean square differentiable. We use polar coordinates for our analysis and the path loss function is then given as:

$$l(r) = \begin{cases} 1 & \text{for } r \leq r_0, \\ f(r) & \text{for } r > r_0. \end{cases} \tag{3.48}$$

The evaluation of the differential of the covariance kernel $c(t)$ for $t \in [0, \delta]$ is achieved by splitting the integration. Since the covariance kernel is an even function, the differential function $c^{'}(t)$ is an odd function. Thus, if the value of the differentiation of the covariance kernel at $t = 0$ is non-zero, the covariance kernel is not differentiable at $t = 0$.

Using polar coordinates, the covariance kernel in (3.7) is given by $c(t) = p^2 \int_0^{2\pi} \int_0^{\infty} l(r) l(\sqrt{r^2 + t^2 - 2rt\cos(\theta)}) r dr d\theta$. Without loss of generality, consider the two points separated by a distance of t to be origin O and $\mathbf{t} = (t, 0)$. Consider two discs of radius r_0 around O and $\mathbf{t}$. Then, one can split the entire plane into six regions as illustrated in Figure 3.10. The covariance kernel is the integration of product of two path loss functions over entire plane. The integrand in R_1, R_2 is equal to one since the

distance from any point in these regions to O and $\mathbf{t}$ is less than r_0. Similarly, it is $f(r)$ and $f(\sqrt{r^2+t^2-2rt\cos(\theta)})$ in R_3 and R_4 respectively and $f(r)f(\sqrt{r^2+t^2-2rt\cos(\theta)})$ in R_5 and R_6.

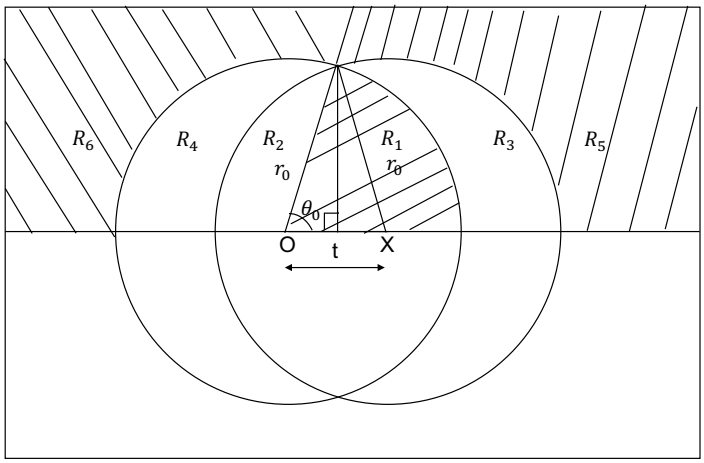


Figure 3.10: Figure illustrating various regions of the integration in the case with path lass function of class-2.

Thus, we have

$$
\begin{aligned}
c(t) = p^2\Bigg[& \int_0^{\theta_0}\int_0^{r_0} r dr d\theta + \int_{\theta_0}^{\pi}\int_0^{t\cos\theta+\sqrt{r_0^2-t^2\sin^2\theta}} r dr d\theta + \\
& \int_0^{\theta_0}\int_{r_0}^{t\cos\theta+\sqrt{r_0^2-t^2\sin^2\theta}} f(r) r dr d\theta + \\
& \int_{\theta_0}^{\pi}\int_{t\cos\theta+\sqrt{r_0^2-t^2\sin^2\theta}}^{r_0} f(\sqrt{r^2+t^2-2rt\cos(\theta)}) r dr d\theta + \\
& \int_0^{\theta_0}\int_{t\cos\theta+\sqrt{r_0^2-t^2\sin^2\theta}}^{\infty} f(r) f(\sqrt{r^2+t^2-2rt\cos(\theta)}) r dr d\theta \\
& + \int_{\theta_0}^{\pi}\int_{r_0}^{\infty} f(r) f(\sqrt{r^2+t^2-2rt\cos(\theta)}) r dr d\theta \Bigg].
\end{aligned}
\tag{3.49}
$$

Now, we differentiate the covariance kernel by exchanging the differ-

entiation and the integration in each and every term of the sum. Then, the differential of the covariance kernel at $t = 0$ is given by

$$c^{'}(0) = p^2[-2r_0 + 2r_0 f(r_0) - r_0 f(r_0) f(r_0)]. \tag{3.50}$$

Thus, from (3.50), we get that if the path loss function given by (3.47), belongs to Class-2 i.e., if its continuous, or equivalently if $f(r_0) = 1$, then the limiting Gaussian field is mean-square differentiable.

Similar analysis can be done for path loss functions which belong to Class-3 as given in (3.26). The area of integration in this case is illustrated in Fig.3.11. Thus, the covariance function is given as

$$c(t) = p^2[2r_0(2r_0 - t)],$$

and $c'(0) = -2r_0$. Therefore, the limiting Gaussian field is not mean-square differentiable and also not sample path differentiable since it has no L^2 derivatives [64].

3.7.6 Evaluation of variance

Grid Users. Let $R^g_{\lambda_b} = \sum_{i=1}^{m} \mathcal{S}_{\lambda_b}(\mathbf{y}_j)$, where m is the number of users in the region A, and $\mathbf{y}_i's$ are the grid locations in A. Then, the variance can be evaluated as:

$$\begin{aligned} \text{var}[R^g_{\lambda_b}] &= \text{var}[\sum_{i=1}^{m} \mathcal{S}_{\lambda_b}(\mathbf{y}_j)] = m\text{var}[\mathcal{S}_{\lambda_b}(\mathbf{y}_j)] + \sum_{i,j=1, i\neq j}^{m} \text{cov}[\mathcal{S}_{\lambda_b}(\mathbf{y}_i), \mathcal{S}_{\lambda_b}(\mathbf{y}_j)] \\ &= \frac{\mu^2 m}{\kappa^2 \lambda_b} c(0) + \frac{\mu^2}{\kappa^2 \lambda_b} \sum_{i,j=1, i\neq j}^{m} c(||\mathbf{y}_i - \mathbf{y}_j||) \end{aligned} \tag{3.51}$$

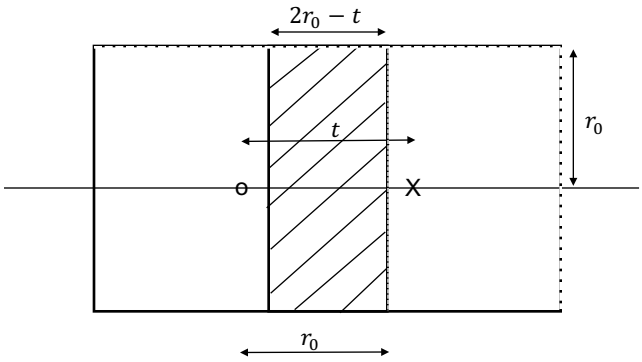


Figure 3.11: Figure illustrating the region of integration in the case with path loss function of class-3.

Random Users. Let $R^r_{\lambda_b} = \sum_{i=1}^{N} \mathcal{S}_{\lambda_b}(\mathbf{y}_j)$, where N is Poisson random variable with parameter $\lambda_b|A_{\lambda_b}|$, and $\mathbf{y}_i's$ are uniformly distributed locations in the region A_{λ_b}. Also, $\mathcal{S}_{\lambda_b}(\mathbf{y}_j)$ are considered to be Gaussian random variables with mean μ and variance $\frac{\mu^2 c(0)}{\kappa^2 \lambda_b}$. Then, the variance can be evaluated using the conditional variance formula:

$$\begin{aligned}
\text{var}[R^r_{\lambda_b}] &= \text{var}[\mathbb{E}[R^r_{\lambda_b}|N=n]] + \mathbb{E}[\text{var}[R^r_{\lambda_b}|N=n]] \\
&= \text{var}[N\mu] + \mathbb{E}[N\big(\frac{\mu^2 c(0)}{\kappa^2\lambda_b}\big)] + \mathbb{E}[\sum_{i,j=1,i\neq j}^{n} \text{cov}\Big[\mathcal{S}_{\lambda_b}(y_i), \mathcal{S}_{\lambda_b}(y_j)|N=n\Big] \\
&= \mu^2\lambda_u|A_{\lambda_b}| + \lambda_u|A_{\lambda_b}|\frac{\mu^2 c(0)}{\kappa^2\lambda_b} + \\
&\quad \mathbb{E}\Big[N(N-1)\frac{1}{|A_{\lambda_b}|^2}\int_{A_{\lambda_b}}\int_{A_{\lambda_b}} \text{cov}[\mathcal{S}_{\lambda_b}(y_i), \mathcal{S}_{\lambda_b}(y_j)]d\mathbf{x}d\mathbf{y}\Big] \\
&= \mu^2\lambda_u|A_{\lambda_b}| + \lambda_u|A_{\lambda_b}|\frac{\mu^2 c(0)}{\kappa^2\lambda_b} + \lambda_u^2\Big[\frac{\mu^2}{\kappa^2\lambda_b}\int_{A_{\lambda_b}}\int_{A_{\lambda_b}} c(||\mathbf{x}-\mathbf{y}||)d\mathbf{x}d\mathbf{y}\Big].
\end{aligned} \tag{3.52}$$

Chapter 4

Mobility-driven Association Policies

4.1 Introduction and Related Work

Modeling and improving handover performance has been extensively addressed in the cellular network literature. Emphasis was on offloading the mobile user traffic to macro-cells in heterogeneous networks to reduce the frequency of handovers, but the increase in the volume of mobile users will lead to heavy congestion at macro-cells.

A handover management technique, based on self-organizing maps, is proposed in [65] to reduce unnecessary handovers for indoor users in two-tier cellular networks. The authors in [66] present a study to avoid unnecessary vertical handovers and reduce the overall packet delay for low-speed users in two-tier downlink cellular networks. Handover signaling overhead reduction algorithms are proposed in [67] for two-tier networks and in [68] for cloud-RAN based heterogeneous networks. Handover delay for a single tier network is characterized in [69]. However, none of the aforementioned studies tackle the interplay between handover cost and capacity gain as a function of the base station intensity.

In this chapter, we explore a class of mobility-driven association poli-

cies that optimize the trade-off between handover cost and throughput gains. There have been many studies in user association policies other than the closest BS policy for different spatial models such as single tier (homogeneous) and multi-tier (heterogeneous) networks, see [70], [71], [72], [73], [74], [75], [76] and references therein. Primarily they can be classified into three categories based on the metrics considered for association: (1) received signal strength, i.e., highest SINR (or highest SIR) in interference limited networks (2) load balancing (for heterogeneous networks, it is often coupled with inter-cell interference mitigation [77]) and, (3) energy aware, where one considers the operational power consumption of base stations as the user association metric so that they are likely to be associated with energy-efficient base stations [78].

Most of these works accurately model spatial traffic variations but none consider user mobility. Although, in the context of multi-hop ad-hoc networks, a selection of a receiver for relaying the information that is aware of the direction of the destination is studied in [79].

We model the cellular network using stochastic geometry, which is a widely accepted mathematical tool to model and analyze cellular networks which enables performance characterization in terms of the base station intensity, λ as well as other physical layer parameters (see [6] for a survey). In such networks, under the closest base station association policy, the intensity of handovers scales very poorly, i.e., proportionally to $\sqrt{\lambda}$ [80]. Thus, our proposed association policies are also geared at reducing the scaling of the intensity of handovers.

The handover rate for Poisson cellular networks in terms of the base station intensity is characterized in [80] for single-tier scenario and in [81] for multi-tier scenarios. However, none of the aforementioned studies investigate the integrated effect of network densification (i.e., BS intensity) in terms of both the handover cost and the throughput gains. The authors in [82], [83] proposed a simple handover skipping scheme to mitigate handover cost in single tier networks, where they advocate sacrificing the best base station association to skip some handovers along the user trajectory. Although the proposed skipping scheme improves user throughput at higher velocities, it is topology agnostic, which can result in non-efficient skipping decisions. In [82], the authors exploited topology awareness and user trajectory estimation to propose smart handover management schemes in single and two-tier downlink cellular networks. However, it is difficult to conduct tractable analysis for the proposed handover skipping schemes due to the random shape of the Voronoi cell and the random location and orientation of the trajectory within the Voronoi cell.

4.2 System Model

Consider an ultra-dense wireless network with base stations denoted through their locations on the Euclidean plane. The configuration of the base stations, $\Phi = \{B_1, B_2, ..\}$, is assumed to be a realization of Poisson point process of intensity λ in $\mathbb{R}^2$. Often, we let B_i denote both the location of the base station and a state representing the users' association with base station

B_i. We consider downlink transmissions and assume that all base stations transmit at a fixed power ξ. We shall adopt a deterministic isotropic bounded path loss model represented by a function $l(\mathbf{.})$.

We consider a tagged mobile user moving at a constant velocity v along a trajectory, T, starting from the origin o and moving to a destination d. We assume that the user is aware of the base station locations, denoted as $\phi = \{b_1, b_2, ...\}$.

In general, handovers are performed in three phases: initiation, preparation, and execution [84]. This involves signaling overhead between the user, the associated base station, the target base station and core network, which often interrupts the data flow to the user. We denote the duration of each interruption (handover delay) by δ seconds, measured from the beginning of the initiation phase to the end of the execution phase. Such delays can be significant [85]. In the sequel, we model the cost of a handover by considering the loss in throughput associated with such handover delays. Further, we also introduce an additional fixed loss of c bits per handover to model the packet loss and signaling overhead.

4.3 Mobility-driven Association Policies

In this section, we introduce the notion of mobility-driven association policies which represents feasible associations encountered by a typical tagged mobile user.

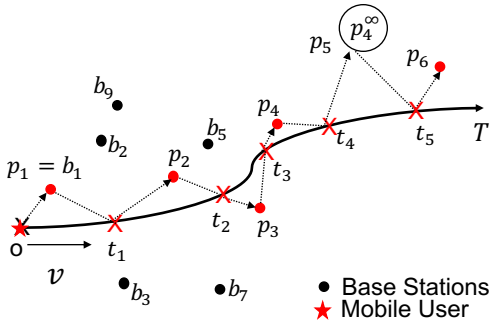


Figure 4.1: Different associations states a tagged mobile user will be in while moving along its trajectory T at a constant velocity v.

The tagged user moving along its trajectory goes through different *association states*, p_i that are either a connected (associated base station) or disconnected state as illustrated in Figure 4.1. Let $\mathbf{p} = (p_1, p_2, ...p_k)$ denote the sequence of such association states of length $k \in \mathbb{N}$ and $\mathcal{P}$ denote set of feasible association sequences.

Disconnected States. We shall distinguish various disconnected states for a purpose that will be clear in the next section. If the user is associated with base station b_i and becomes disconnected, we denote it as being in state b_i^∞. Similarly, if the user is disconnected at the beginning of motion, i.e., at the origin o, we denote it as being in state b_o^∞. Let ψ denote the set of possible disconnected states.

Given an association sequence, $\mathbf{p} \in \mathcal{P}$ of length k, one can choose a sequence of handover times $\mathbf{t} = (t_1, t_2, \ldots, t_{k-1})$, where t_i denotes the handover time from $p_i \rightarrow p_{i+1}$. Let $\mathcal{T}_\mathbf{p}$ denote the set of all such feasible sequences of

handover times for $\mathbf{p}$.

The feasibility of an association sequence depends on the tagged mobile user's:

- *Trajectory.* We assume that user is moving along a straight line. Without loss of generality, we assume $T = \{u(t) = vt : t \in [0, t(d)]\}$, where $u(t)$ denotes the user's location on the trajectory at time t and $t(u)$ denotes the time at which the tagged user reaches the location u on the trajectory.

- *Possible associations.* We assume that the mobile user can only connect to a base station from a set of *accessible base stations.* The mobile user is assumed to be disconnected in the case where there are no accessible base stations. Further, we assume that if the mobile user can be connected, it will be connected, i.e., the user is disconnected only in the absence of accessible base stations. This is in part inspired by the threshold service model considered in [59], where the user is idle in the case where the SNR received is less than a given threshold.

Accessible base stations. For the mobile user at location u on T, we define the set of accessible base stations, $\Phi_\gamma(u)$, parameterized by γ, as

$$\Phi_\gamma(u) = \Phi \cap A_\gamma(u), \tag{4.1}$$

where the association region $A_\gamma(u)$ is a convex association region centered at u and parameterized by γ. The parameter γ can either be a scalar, e.g., a disc

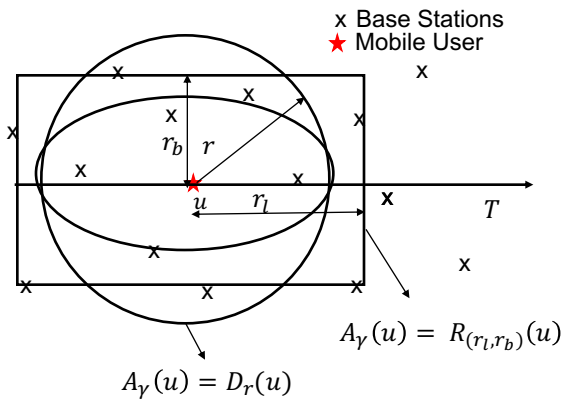


Figure 4.2: Different possible association regions for a mobile user at u, $A_\gamma(u)$.

of radius, r, $A_\gamma(u) = D_r(u)$ or a vector, $(2r_l, 2r_b)$ representing side lengths of a rectangle $A_\gamma(u) = R_{(r_l,r_b)}(u)$ centered at u, or parameters for an ellipse as illustrated in Figure 4.2.

Maximal Connection Intervals. We define the longest connection interval for a base station, b_i, as the longest line segment on the user's trajectory on which the tagged user can associate with base station b_i, denoted by $\left[u_{b_i}^{(1)}, u_{b_i}^{(2)}\right]$, where $u_{b_i}^{(1)}, u_{b_i}^{(2)}$ are locations on the user's trajectory, T. Figure 4.3 illustrates the interval for an association region, $D_r(u)$, denoting a disc of radius r, in which case, for $D_r(b_i)$ namely a disc of radius r centered at b_i, $u_{b_i}^{(1)}, u_{b_i}^{(2)}$ will be the intersection of this disc with the user's trajectory T.

Handover Sets. Let us define the handover set, $\mathcal{B}(b_i)$ as the set of base stations to which the mobile user could realize a handover from base station $b_i \in \phi$, i.e., all the base stations that are accessible from some point on the maximal connection interval of b_i, $\left[u_{b_i}^{(1)}, u_{b_i}^{(2)}\right]$. Further, for any disconnected

state b_i^∞ the handover set $\mathcal{B}(b_i^\infty)$ includes only the first available base station along the trajectory:

$$\mathcal{B}(b_i^\infty) = \left\{ b_j \in \phi \;\; \text{s.t} \;\; u_{b_j}^{(1)} = \min_{b_k \in \phi} \left[u_{b_k}^{(1)} : u_{b_k}^{(1)} > u_{b_i}^{(2)} \right] \right\}. \tag{4.2}$$

Given the properties of the Poisson point process, this is a singleton almost surely. Note that the mobile user enters the disconnected state from a base station b_i only when $\mathcal{B}(b_i)$ is empty.

Recall that $\phi_r(o), \phi_r(d)$ denote the set of accessible base stations for the tagged user at the origin o and the destination d respectively. Then, for given trajectory and possible associations for the tagged user, we now formally define feasible association sequences.

Definition 9. *A feasible association sequence, $\boldsymbol{p} = (p_1, p_2, \ldots, p_k)$, of length $k \in \mathbb{N}$, for the trajectory T is such that:*

- *p_1 is either a base station in $\phi_r(o)$ or the disconnected state b_o^∞, when $\phi_r(o)$ is empty,*
- *for each $i > 1$, p_i is either a base station in $\mathcal{B}(p_{i-1})$ or the disconnected state p_{i-1}^∞, when $\mathcal{B}(p_{i-1})$ is empty, and*
- *$p_k \in \phi_r(d)$.*

Thus, a specific mobility-driven association policy is equivalent to selecting a pair $(\mathbf{p}, \mathbf{t})$, with $\mathbf{p} \in \mathcal{P}$ and $\mathbf{t} \in \mathcal{T}_\mathbf{p}$. We are interested in finding a

policy that optimizes the *reward* that captures both the performance seen by the user and the cost of handovers as defined below.

Reward. For each pair $(\mathbf{p}, \mathbf{t})$ with $\mathbf{p} \in \mathcal{P}$ and $\mathbf{t} \in \mathcal{T}_{\mathbf{p}}$, we define the total volume of data delivered to the tagged mobile user during its trajectory T as the reward, $r(\mathbf{p}, \mathbf{t})$.

For $p_i \in \mathbf{p}$, we define the Shannon rate, $s_{p_i}(t)$, seen by the user at any given time t as:

$$s_{p_i}(t) = \omega \ln \left(1 + \frac{\xi \times l(||p_i - u(t)||)}{\sigma^2}\right), \text{ for } t_{i-1} \leq t < t_i, \tag{4.3}$$

where σ^2 is the noise power, ω is the system bandwidth and $||.||$ denotes Euclidean distance. Since the rate seen by the user in the disconnected state is zero, we set $s_{p_i}(t) = 0$, if $p_i \in \psi$.

Thus, for an association sequence, $\mathbf{p} = (p_1, p_2, \ldots, p_k)$, and associated sequence of handover times, $\mathbf{t} = (t_1, t_2, \ldots, t_{k-1})$, the total volume of data delivered, $r(\mathbf{p}, \mathbf{t})$, is the sum of the data volumes delivered by individual base stations, denoted by $r(p_i, (t_{i-1}, t_i))$, i.e., $r(\mathbf{p}, \mathbf{t}) = \sum_{i=1}^{k} r(p_i, (t_{i-1}, t_i))$, where $t_0 = 0, t_k = t(d)$ and

$$r(p_i, (t_{i-1}, t_i)) = \left[\int_{t_{i-1}+\delta/2}^{t_i-\delta/2} s_{p_i}(t)dt - c\right]^+. \tag{4.4}$$

We model the cost of a handover by assuming that data transmission is interrupted for δ seconds, equally split between the two states involved in handover. Further, the additional fixed loss of c bits in the data per handover

Symbol	Definition
$\phi = \{b_1, b_2, \dots\}$	Set of base stations
T	Mobile user trajectory of motion
$\mathcal{B}(b_i)$	Handover set of base station b_i
$\left[u_{b_i}^{(1)}, u_{b_i}^{(2)}\right]$	Interval during which the user can associate with base station b_i
$\mathbf{p} = (p_1, p_2, \dots, p_k)$	Association Sequence of length k
t_i	Handover time from p_i to p_{i+1}
$\mathbf{t} = (t_1, t_2, \dots, t_{k-1}),$	Handover times for association path $\mathbf{p}$
$r(p_i, p_{i+1})$	Net rate associated with the handover $p_i \rightarrow p_{i+1}$
δ, c	Time of service interruption due to a handover and a fixed cost

Table 4.1: Table of Notation for the Dynamic Programming.

is modeled by subtracting c bits from the total volume of data delivered from each base station. This captures packet losses and signal overheard.

In cases where, (1) the value of the upper limit is less than the lower limit, i.e., when the time taken for handover is more than the users association time or (2) the volume of data delivered is less than the penalty c, we model the total volume of data delivered to be zero.

4.3.1 Optimal Association Policy

Problem 1. In this setting, the optimal handover problem consists in finding a sequence of base stations and handover times that results in the maximum volume of data, i.e.,

$$r^* = \max_{\mathbf{p},\mathbf{t}} \{r(\mathbf{p}, \mathbf{t}) : \mathbf{p} \in \mathcal{P}, \mathbf{t} \in \mathcal{T}_{\mathbf{p}}\}, \tag{4.5}$$

where $\mathcal{P}, \mathcal{T}_{\mathbf{p}}$ denote the set of all possible association sequences and associated handover times respectively.

Recall that $\left[u_{b_i}^{(1)}, u_{b_i}^{(2)}\right]$ is the maximal connection interval for base station b_i. Given that if the tagged user can be connected, it will be connected, a handover from p_i to the disconnected state happens at location $u_{p_i}^{(2)}$. Similarly, a handover from the disconnected state to base station p_{i+1} happens at location $u_{p_{i+1}}^{(1)}$.

For each handover from base station $p_i \to p_{i+1}$, we have a continuum of possible handover times with some constraints that we will not specify. Thus, the above problem is a mixed optimization with discrete and continuous variables can be very challenging to solve. Further, the handover time t_i is functionally dependent on the entire association sequence.

Let us discretize the time with step size δ and consider a decision tree with origin o as the parent node and feasible association base stations and disconnected states at each time step as a new layer in a decision tree. If m denotes the total number of time steps, then the number of nodes in such a tree is in the order of $(2\lambda(\pi r^2))^m$. Further, the combinatorial optimization problem reduces to finding a maximum reward path in this decision tree, which has a complexity of $O((2\lambda(\pi r^2))^m)$ [86].

In next section, we will restrict the set of feasible association sequences by considering *forward association sequences* and defining a single sequence of handover times for each association sequence. We then propose an algo-

rithmic solution with less complexity to the problem of finding the optimal policy, specifically a dynamic programming approach. We do not claim to solve Problem 1.

4.4 Forward Association Policies

Let us first, define the projection of a base station location, $b_i \in \phi$ on the trajectory, T as $\Pi_T(b_i)$. Given the set of base stations is a realization of a homogenous Poisson point process, we have that all projections of the base station locations on the straight line trajectory are almost surely unique, i.e., no two projections are the same.

We then define forward association sequences based on forward handover sets:

Definition 10. *The forward handover set, $\mathcal{B}^f(b_i) \subset \mathcal{B}(b_i)$ is the set of base stations to which the mobile user could realize a handover from base station b_i that are:*

1. *within a distance r from some point on the interval $\left[\Pi_T(b_i), u_{b_i}^{(2)}\right]$ and such that*
2. *for any $b_j \in \mathcal{B}^f(b_i)$, we have $\Pi_T(b_j) > \Pi_T(b_i)$.*

The base stations in $\mathcal{B}^f(b_i)$ are in the region illustrated in Figure 4.3. Note that the definition for the handover set corresponding to the disconnected state b_i^{∞} is the same as before (4.2). Now, a forward association sequence,

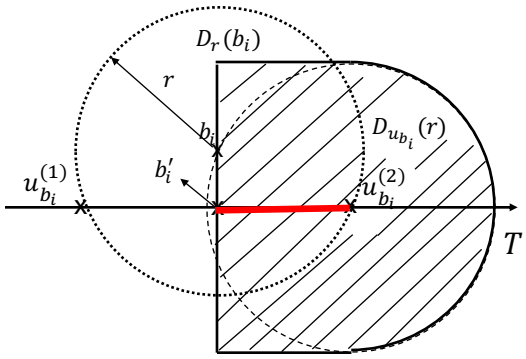


Figure 4.3: Figure illustrating the handover sets for an association region of a disc of radius r denoted as $D_r(u)$.

$\mathbf{p} = (p_1, p_2, ..., p_k)$ is an association sequence where for every, $i > 1$, p_i is either a base station in $\mathcal{B}^f(p_{i-1})$ or the disconnected state p_{i-1}^{∞}. Let $\mathcal{P}_f \subset \mathcal{P}$ denote the set of all possible forward association sequences.

Intuitively forward association sequences result in fewer handovers and longer connection periods. Let us now define a forward association graph as follows:

Definition 11. *For a given trajectory, T, the set of base stations, ϕ, and possible disconnected states, ψ, we define a forward association graph $G = (V, E)$, where $V = \phi \cup \psi \cup \{o, d\}$ and $E \subset V \times V$. An edge $(p_i, p_{i+1}) \in E$, for all pairs (p_i, p_{i+1}) of forward association sequences in $\mathcal{P}_f$. Similarly, edges $(o, p_1), (p_k, d) \in E$, for all p_1, p_k of forward association sequences in $\boldsymbol{p} \in \mathcal{P}_f$.*

The forward association graph G is a directed acyclic graph. This is due to: (1) the constraint of the forward handover sets and (1) the different disconnected states based on the last association base station. Further, the

forward association sequences are the paths in the graph from node o to node d. Figure 4.4 gives an instance of such a graph.

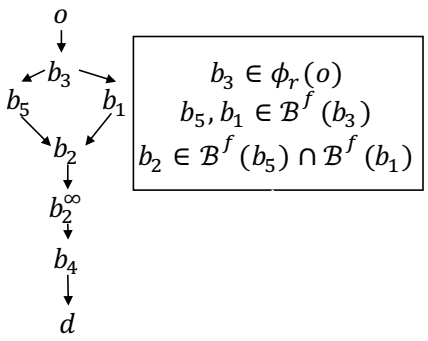


Figure 4.4: Figure illustrating the directed acyclic graph.

In the sequel, we will evaluate the reward, $r(\mathbf{p}, \mathbf{t})$ associated with the forward association path $\mathbf{p}$ by representing it as the sum of rewards on the edges along the path. Thus, we further restrict ourselves to the case where the sequence of handover times, $\mathbf{t}$, is a function of the association sequence $\mathbf{p}$ by selecting the best handover time from the continuum under the constraint that $t_i \in \mathbf{t}$ depends only on the pair $(p_i, p_{i+1}) \in \mathbf{p}$.

Note that, for a handover from base station p_i to p_{i+1}, it is best to handover at a time where the mobile is equidistant to both base stations. Let $u_{p_i,p_{i+1}}$ be the location on the trajectory T where the user is equidistant from both the base stations, i.e., $||p_i - u_{p_i,p_{i+1}}|| = ||u_{p_i,p_{i+1}} - p_{i+1}||$. We assume that the handover time t_i for handover $p_i \rightarrow p_{i+1}$, for $p_i, p_{i+1} \in \phi$, to be within the range $\left[t(p_i^{'}), t(p_{i+1}^{'})\right]$. In the event the ideal time, $t(u_{p_i,p_{i+1}})$ is not in this

range, we handover at one of the extreme points of the interval $\left[t(p_i^{'}), t(p_{i+1}^{'})\right]$. More precisely, we make the following assumption:

Assumption 1. We assume that the handover time from p_i to p_{i+1} is given by:

$$t_i = \begin{cases} t(p_i^{'}) \;\text{ if }\; t(u_{p_i,p_{i+1}}) < t(p_i^{'}) \;\text{ for }\; p_i \to p_{i+1}, \\ t(p_{i+1}^{'}) \;\text{ if }\; t(u_{p_i,p_{i+1}}) > t(p_{i+1}^{'}) \;\text{ for }\; p_i \to p_{i+1}, \\ t(u(p_i,p_{i+1})) \;\text{ if }\; t(u_{p_i,p_{i+1}}) \in [t(p_i^{'}), t(p_{i+1}^{'})] \;\text{ for }\; p_i \to p_{i+1}, \\ t(u_{p_i}^{(2)}) \;\text{ for }\; p_i \to b_j^{\infty}, \\ t(u_{p_{i+1}}^{(1)}) \;\text{ for }\; b_j^{\infty} \to p_{i+1}. \end{cases} \tag{4.6}$$

Figure 4.5 illustrates all cases of handover times for a simple example of mobile user moving at unit velocity.

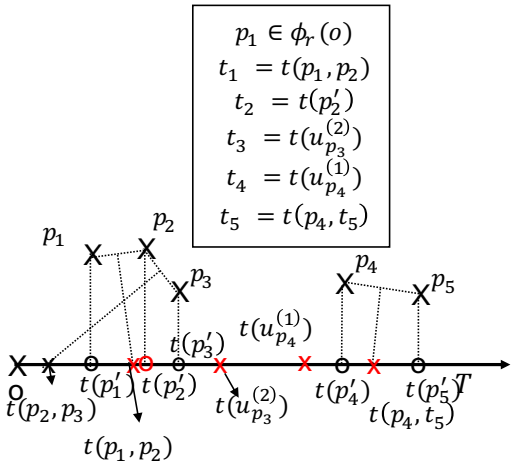


Figure 4.5: Figure illustrating all cases of handover times for a mobile user moving along a straight line at unit velocity.

Having defined handover times as a function of the forward association sequence, the association reward, $r(\mathbf{p}, \mathbf{t})$ now depends only on the association

sequence and is denoted by $r(\mathbf{p})$. Let $r(p_i, p_{i+1})$ denote the total volume of data received by the mobile user while moving from $p_i^{'} \rightarrow p_{i+1}^{'}$ under Assumption 1. Then, the association reward, $r(\mathbf{p})$ for a forward association sequence of length k, is additive, and with $p_0 = o, p_{k+1} = d$, we have $r(\mathbf{p}) = \sum_{i=0}^{k} r(p_i, p_{i+1})$, where,

$$r(p_i, p_{i+1}) = \left[\left(\int_{t(p_i^{'})+\delta/2}^{t_i-\delta/2} s_{p_i}(t)dt + \int_{t_i+\delta/2}^{t(p_{i+1}^{'})-\delta/2} s_{p_{i+1}}(t)dt \right) - c \right]^{+}, \quad (4.7)$$

with $s_{p_i}(t)$ as in (4.3) and the cost of handover modeled as before.

4.4.1 Optimal Forward Association Policy

Problem 2. In this setting, the problem of finding the optimal sequence of base stations and handover times that results in the maximum total data delivered to the tagged user under the constraints of forward association and handover sequences is equivalent to solving the largest weight path from node o to node d [87], i.e.,

$$r_f^* = \max_{\mathbf{p}} \{ r(\mathbf{p}) : \mathbf{p} \in \mathcal{P}_f \}, \quad (4.8)$$

where, r_f^* denotes the optimal reward and $\mathcal{P}_f$ denotes the set of all forward association sequences, i.e., paths from o to d.

Given that ϕ is a realization of a Poisson point process of intensity λ, the mean number of base stations seen by the tagged mobile user along its trajectory that are within a distance r from its path is equal to $2rd\lambda$, where d is the total distance traversed by the user. The mean number of nodes in

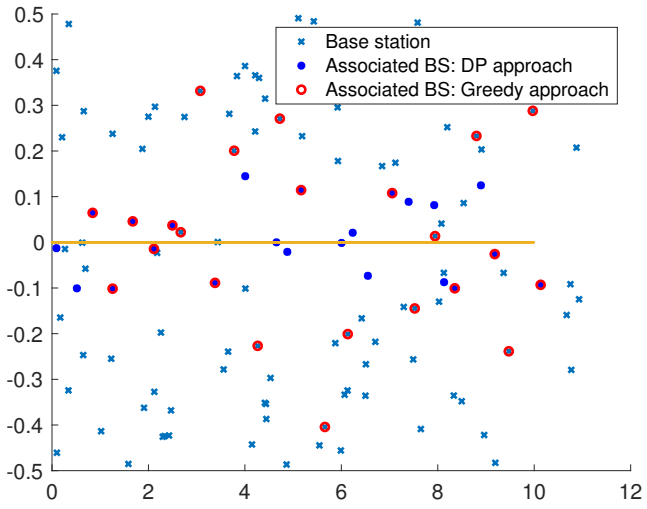


Figure 4.6: Figure illustrating the optimal association path given by the dynamic programming approach.

the directed acyclic graph is of the order of the number of base stations, i.e., $|V| = O(2rd\lambda)$. Now, a given base station can have edges to base stations in the region as illustrated in Fig 4.3, thus we have that $|E| = O(\lambda^2 r^3 d)$. Thus, the mean time complexity of solving the optimal path in the graph is $O(|V| + |E|) = O(\lambda^2 r^3 d)$ [86].

Thus, Problem 2 can be solved using dynamic programming with a linear time complexity with respect to the total distance traversed. Considering a certain realization of base stations, we solve the problem using Dijkstra's algorithm ([86]). Figure 4.6 illustrates the optimal base stations as given by dynamic programming along with the closest base station. Further, the performance analysis of this approach is studied in Section IV.

Symbol	Definition
$\Phi_\gamma(u)$	Accessible base station for mobile user at u
$(P(t), t > 0)$	Random process denoting the association base station to the mobile
$(D(t), t > 0)$	Random process denoting the distance to the association base station
$(S(t), t > 0)$	Random process denoting Shannon rate seen by the mobile user
$(T_i)_{i=1}^{\infty}$,$u(T_i)_{i=1}^{\infty}$	Handover times and the location of the mobile user
$\mathbf{Z}_n = (X_n, Y_n)$	Relative position of the association base station w.r.t. mobile user's location $u(T_n)$
ρ	Rate of handovers
$\mathcal{S}(\gamma)$	Effective throughput

Table 4.2: Table of Notation for Greedy Association.

Thus, with knowledge of the entire layout of base stations along the mobile user's trajectory, one can find the optimal set of base station associations and handover times using dynamic programming. In the sequel, we assume limited information about the base stations in a specific environment and propose a greedy association policy.

4.5 Farthest Greedy Policy

In this section, we consider networks with high handover costs and propose association policies where the mobile user greedily associates with the base station that maximizes the connection time, i.e., the base station that is farthest out in the direction of its motion. Handovers are constrained to a fixed region around the mobile, which represents both a constraint on the

information available as well as the degree of greediness in avoiding handovers. In this setting, as the first step, we analyze these policies using Markov chains. Since associating with a base station that is at a large distance may result in poor throughput, we then optimize the performance of such greedy association policies with respect to the size of the association region in the second step.

We still assume that the base stations, Φ, are a realization of a Poisson point process of intensity λ and that the tagged user is moving along the x-axis, starting from the origin, i.e., $T = \{u(t), t \geq 0\}$. As already mentioned, to evaluate the performance of greedy policies, we establish a connection with the theory of Markov processes. The steady state of these processes is analyzed. Thus, we assume that the tagged users' trajectory has infinite length.

We model the constraints on possible handovers and the information by assuming that, at any location u on T, the tagged mobile user is only aware of the set of accessible base stations $\Phi_\gamma(u)$ that are within the association region $A_\gamma(u)$ as defined in (4.1).

If we assume a single disconnected state denoted by b^∞, then at $t = 0$, the mobile user is either in the disconnected state b^∞ or connected to a base station in $\Phi_\gamma(o)$. Given that at this time, the tagged user is not aware of base station locations beyond the region $A_\gamma(o)$, and that they are randomly located according to a Poisson point process, the first handover time T_1 is a random variable. We retain the definitions of maximal connection intervals, $\left[u_{B_i}^{(1)}, u_{B_i}^{(2)}\right]$ and handover sets $\mathcal{B}(B_i)$ for a base station $B_i \in \Phi$ as in the previous sections.

Definition 12. *An association process,* $(P(t), t \geq 0)$*, is a discontinuous stochastic process denoting the association state of the tagged mobile user at time* t*, with discontinuities at* $(T_i)_{i=1}^{\infty}$ *denoting the handover times, given by*

$$P(t) \in \begin{cases} \{\Phi_\gamma(o), b^\infty\}, & \text{for } \ 0 \leq t < T_1, \\ \{\mathcal{B}(P(T_{i-1})), b^\infty\}, & \text{for } \ T_{i-1} \leq t < T_i. \end{cases} \tag{4.9}$$

Recall that $\Pi_T(B_i)$ denotes the projection of the base station's location on the trajectory T. We define the farthest greedy policy as follows:

Definition 13. *In the farthest greedy policy, when a handover is required, the mobile user chooses as next station the accessible base station that is farthest out in the direction of its motion with forward progress, i.e., for* $T_0 = 0$ *and* $n \geq 1$*,*

$$P(T_n) = \underset{B_i \in \Phi_\gamma(u(T_n))}{\arg\max} \left[||\Pi_T(B_i) - u(T_n)|| : \Pi_T(B_i) > \Pi_T\left(P(T_{n-1})\right) \right], \tag{4.10}$$

and stays connected to that base station for the maximum amount of time, i.e.,

$$T_{n+1} = t\left(u^{(2)}_{P(T_n)}\right). \tag{4.11}$$

Reward. In the sequel, we use *effective throughput* as the metric for the user's performance, which encapsulates both the cost of handovers and the Shannon rate seen by the user. Let us first define the distance process $(D(t), t \geq 0)$ denoting the distance from the mobile user to its association base station when connected and infinity otherwise as given by:

$$D(t) = \begin{cases} \infty \text{ for } P(t) = b^\infty, \\ ||u(t) - P(t)|| \ \text{ otherwise,} \end{cases} \tag{4.12}$$

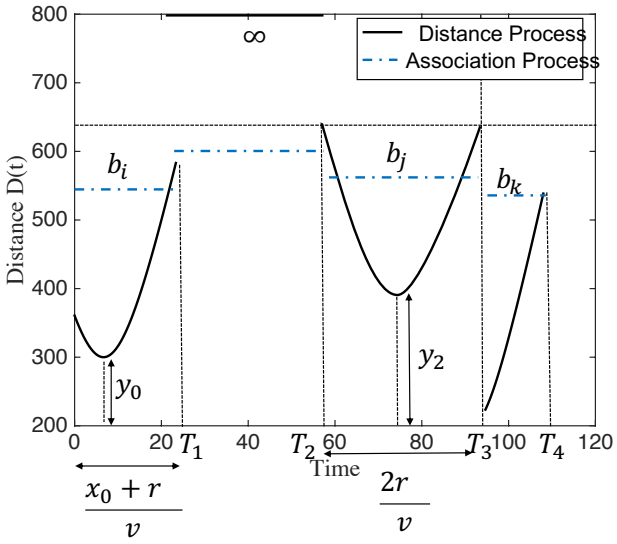


Figure 4.7: Distance from the association base station to the mobile user.

where $||.||$ denotes Euclidean distance. Figure 4 illustrates a realization of association and distance processes. Then, considering SNR, the Shannon rate process, $(S(t), t \geq 0)$ is given by:

$$S(t) = \omega \ln \left(1 + \frac{\xi \times l(D(t))}{\sigma^2}\right). \tag{4.13}$$

Thus, by introducing a cost of a handover similar to that in our previous analysis, we have the following definition:

Definition 14. *The effective throughput $\mathcal{S}(\gamma)$ seen by a mobile user is defined as the time average of the Shannon rate seen by the user when its service is interrupted for δ seconds and penalized c bits for every handover.*

Under our proposed association policy, we establish a connection between the random geometry of base stations and the theory of Markov process

with continuous state space. Using this, we give an analytical expression for the stationary effective throughput.

For simplicity, in the sequel, we consider a fixed rectangular shape, $A_\gamma(u) = R_{(r_l,r_b)}(u)$ centered at the tagged users location u, and optimize its size.

4.5.1 Markov Formulation

To begin with, we have the following definition:

Definition 15. *At any given handover time, T_n, we define the relative position as the coordinates, $\boldsymbol{Z}_n = (X_n, Y_n)$ of the association base station, $P(T_n)$ with respect to the origin shifted to the location $u(T_n)$ on the tagged users' trajectory T, if the user is connected and infinity otherwise as given by:*

$$X_n = \begin{cases} ||\Pi_T(P(T_n)) - u(T_n)||, & \text{if } P(T_n) \in \Phi, \\ \infty & \text{otherwise}, \end{cases} \tag{4.14}$$

$$Y_n = \begin{cases} ycoord(P(T_n)), & \text{if } P(T_n) \in \Phi, \\ \infty & \text{otherwise.} \end{cases} \tag{4.15}$$

Let us recall some terminology regarding general state space Markov processes.

Definition 16. *Given a discrete time Markov process, $(\boldsymbol{Z}_n, n \geq 0)$, with $\boldsymbol{Z}_n = (X_n, Y_n)$ with a state space, $\mathcal{Z}$, the transition probability kernel, $\mathbb{P}(\boldsymbol{z}, Q)$ with $\boldsymbol{z} = (w, h)$ and Q a set in the Borel σ algebra of $\mathcal{Z}$ is defined as:*

$$\mathbb{P}(\boldsymbol{z}, Q) = \mathbb{P}(\boldsymbol{z}_n \in Q | \boldsymbol{z}_{n-1} = \boldsymbol{z}). \tag{4.16}$$

Further, for the continuous part of kernel the transition kernel density, $f(\boldsymbol{z}, \bar{y})$ is defined as:

$$\mathbb{P}(\boldsymbol{z}, Q) = \int_{\bar{y} \in Q} f(\boldsymbol{z}, \bar{y}) d\bar{y}. \tag{4.17}$$

Figure 4.8 illustrates the relative position of an association base station.

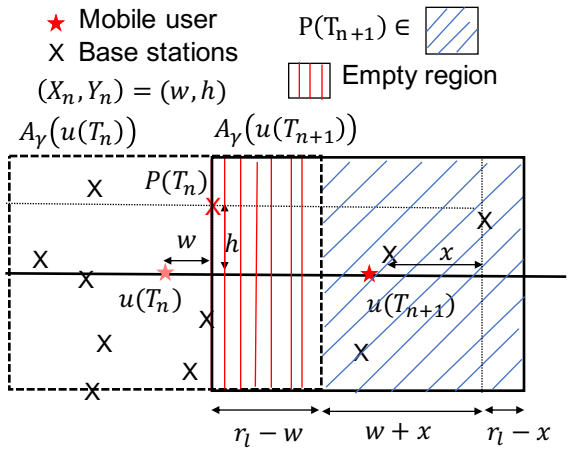


Figure 4.8: Relative position (X_n, Y_n) of the association base station $P(T_n)$ at time instant T_n.

Leveraging the properties of the random geometry of base stations, i.e., the Poisson point process assumptions, we have the following result.

Theorem 16. *Under rectangle-based farthest greedy policy, where the association region is a rectangle $R_{(r_l, r_b)}(u)$ of length $2r_l$ and breadth $2r_b$,*

- *the x-coordinate of the relative position of the association base station $(X_n, n \in \mathbb{N})$ is a Markov process with state space $\mathcal{Z} = \{[-r_l, r_l], \infty\}$ and*

the transition kernel densities for $Q(x) = [-r_l, x]$ *are given by:*

$$f(\infty, Q(x)) = \delta(x - r_l), \mathbb{P}(w, \infty) = e^{-2\lambda r_b(r_l+w)},$$
$$f(w, Q(x)) = \begin{cases} 2\lambda r_b e^{-2\lambda r_b(r_l-x)}, & \text{if } -w \leq x \leq r_l, \\ 0 & \text{otherwise}, \end{cases} \tag{4.18}$$

- *the y-coordinate of the* n^{th} *relative position is given as:*

$$Y_{n+1} = \begin{cases} Y, & \text{if } P(T_{n+1}) \in \Phi, \\ \infty & \text{otherwise}, \end{cases} \tag{4.19}$$

where, $Y \sim U[-r_b, r_b]$ *and the event that* $\{P(T_{n+1}) \in \Phi\}$ *depends only on* X_n*, i.e., given* $X_n = w$*, its probability is given by* $1 - e^{-2\lambda r_b(w+r_l)}$*.*

Proof. Recall that $(T_i)_{i=1}^{\infty}$ denotes handover times and $u(T_i)_{i=1}^{\infty}$ the locations of the mobile user at handover times. The n^{th} handover time, T_n is a stopping time, since the event $\{T_n = t\}$ for $t \geq 0$, depends only on $\mathbf{Z}_1, \mathbf{Z}_2, \ldots, \mathbf{Z}_n$. Then, conditionally on $T_n < \infty$ and $\mathbf{Z}_n = (w, h)$, we have that for all $m \geq 1$:

$$\mathbb{P}(\mathbf{Z}_{n+m} = \mathbf{z} | \mathbf{Z}_i = \mathbf{z}_i \ \forall i < n, \mathbf{Z}_n = (x_n, y_n)) = \mathbb{P}(\mathbf{Z}_{n+m} = \mathbf{z} | \mathbf{Z}_n = (w, h)). \tag{4.20}$$

Recall that Z_{n+1} is the relative position of the association base station, $P(T_{n+1}) \in \Phi \cap R_{(r_l,r_b)}(u(T_{n+1}))$. Given our greedy association policy, we have the following constraint that:

$$P(T_{n+1}) \in \Phi \cap \left[R_{(r_l,r_b)}(u(T_{n+1})) \setminus R_{(r_l,r_b)}(u(T_n)) \right]. \tag{4.21}$$

This is due to the fact that a part of the region, $R_{(r_l,r_b)}(u(T_{n+1}))$ is empty of base stations that is based on its previous association, $P(T_n)$, as illustrated in Figure 4.8.

Further, note that one can determine the location $u(T_n)$ and its corresponding region $R_{(r_l,r_b)}(u(T_n)))$, based on $Z_n = (w, h)$ alone, in particular, only on the x-coordinate of the previous state, i.e., X_n. Thus, the process tracking the x- coordinate of relative position of the association base station $(X_n, n \in \mathbb{N})$ satisfies the Markov property with state space, $\mathcal{Z} = \{[-r_l, r_l], \infty\}$. The y-coordinate Y_{n+1} is independent of Y_n and is uniformly distributed between $[-r_b, r_b]$, given that the user is connected as given in (4.19).

The discrete transition probability to the state "∞", $\mathbb{P}(w, \infty)$ is given by the probability that the region, $R_{(r_l,r_b)}(u(T_{n+1}))\backslash R_{(r_l,r_b)}(u(T_n))$, i.e., $R_2 \cup R_3$ as illustrated in Figure 4.8 is empty. Given that the association region is empty of base stations, the first base station to enter will be at the edge. Thus, the Markov process always transitions from the state "∞" to a state "r_l", and therefore the transition probability kernel, $f(\infty, Q(x))$ will be a Dirac delta function as given in (4.18).

The transition probability kernel for the continuous part is given by the product of the probability that the region R_2, the rectangle of length $(w + x)$ has at least one point and the region R_3, the rectangle of side length $(r_l - x)$ has no points, see Fig 4.8:

$$\mathbb{P}(w, Q(x)) = e^{-\lambda 2 r_b (r_l - x)} \left(1 - e^{-\lambda 2 r_b (w+x)}\right), \quad -w \leq x \leq r_l. \tag{4.22}$$

Thus, the transition kernel density for the continuous and discrete parts of the state space $\mathcal{Z}$ is given by (4.18). □

Let us now focus on proving the existence and uniqueness of its stationary distribution. The Markov process $(X_n, n \geq 0)$ has the following properties:

- *Irreducibility*: The state space of the Markov process, ∞ is an atom. The Markov process is irreducible with respect to this atom.
- *Aperiodic*: There is no such subset of the state space $\mathcal{Z}$, i.e., a partition of the rectangle, such that the probability from transitioning from one subset of the state space (region of the rectangle) to another is one and vice-versa. Thus, the Markov process is aperiodic according to the definition of aperiodicity given in [88].

Given the properties of the Markov process and its transition kernel densities, we have the following result regarding its stationary distribution:

Theorem 17. *Under the rectangle-based association policy, the Markov process, $(X_n, n \geq 0)$ has a stationary distribution with discrete and continuous parts π_d and π_c given by its Laplace transforms:*

$$\begin{aligned}\mathcal{L}_{\pi_c}(s) = &\frac{\left[2\lambda r_b s e^{-sr_l} - 4\lambda^2 r_b^2 e^{-4\lambda r_l r_b + sr_l}\right]\left(\frac{1+\mathcal{L}_{\pi_c}(0)}{2}\right)}{2\lambda r_b s - s^2 - 4\lambda^2 r_b^2 e^{-4\lambda r_l r_b}} \\ &+ \frac{\left[4\lambda^2 r_b^2 e^{-4\lambda r_l r_b - sr_l} - 2\lambda s r_b e^{-4\lambda r_l r_b + sr_l}\right]\left(\frac{1-\mathcal{L}_{\pi_c}(0)}{2}\right)}{2\lambda r_b s - s^2 - 4\lambda^2 r_b^2 e^{-4\lambda r_l r_b}}, \qquad (4.23) \\ \pi_d(r_l) = \pi_d(\infty) = q = &\frac{1 - \mathcal{L}_{\pi_c}(0)}{2},\end{aligned}$$

where,

$$\mathcal{L}_{\pi_c}(0) = \frac{2\lambda r_b\left(1 - e^{2a_1 r_l}\right) + a_1\left(e^{4\lambda r_l r_b} - e^{2a_1 r_l}\right)}{2\lambda r_b\left(1 + e^{2a_1 r_l}\right) - a_1\left(e^{4\lambda r_l r_b} + e^{2a_1 r_l}\right)}, \tag{4.24}$$

with, $a_1 = \lambda r_b + \lambda r_b\sqrt{1 - 4e^{-4\lambda r_l r_b}}$.

Proof. Let us consider the stationary distribution of this process in terms of π_d and π_c associated with the discrete and the continuous parts. Thus, $\pi_c(x)$ represents the continuous part of the stationary distribution of the Markov process for $x \in [-r_l, r_l)$, $\pi_d(r_l)$ and $\pi_d(\infty)$ are the discrete probability masses associated with the states r and ∞. We have the following equations:

$$\begin{aligned}
&\pi_c(x) = 2\lambda r_b e^{-2\lambda r_b(r_l - x)}\left(\int_{-x}^{r_l^-} \pi_c(w)dw + \pi_d(r_l)\right),\\
&\pi_d(\infty) = e^{-2\lambda r_b r_l}\int_{-r_l}^{r_l^-} \pi_c(w)e^{-2\lambda r_b w}dw + \pi_d(r_l)e^{-4\lambda r_b r_l},\\
&\pi_d(r_l) = \pi_d(\infty),
\end{aligned} \tag{4.25}$$

with $\pi_d(r_l) + \pi_d(\infty) + \int_{-r_l}^{r_l^-} \pi_c(w)dw = 1$. We use Laplace transform techniques to solve the integral equations. □

Let the relative position of the association base station conditioned on the availability of an accessible base station be denoted by $(\hat{X}_1, \hat{Y}_1)$. Given the Laplace transform of the stationary distribution, we can evaluate various moments of the conditional random variable $\hat{X}_1$ and also study its asymptotics. The mean of the random variable $\hat{X}_1$ is given by :

$$\mathbb{E}[\hat{X}_1] = -\mathcal{L}'_{\pi_c}(0) = \frac{(1 - e^{-4\lambda r_l r_b})[1 - \mathcal{L}_{\pi_c}(0)] - 4\lambda r_l^2 e^{-4\lambda r_l r_b}}{4\lambda r_l e^{-4\lambda r_l r_b}}. \tag{4.26}$$

In the limit, as the density of base station λ goes to infinity, we have

$$\lim_{\lambda\to\infty} \mathcal{L}_{\pi_c}(0) = 1.$$

Thus, we have that the stationary distribution of the states r_l and ∞ goes to zero, i.e., $\pi_d(r_l) = \pi_d(\infty) = 0$, and for the continuous part we have:

$$\lim_{\lambda\to\infty} \pi_c(x) = \delta(x - r_l),$$

and in the limit the expectation is $\mathbb{E}[\hat{X}_1] = r_l$.

4.5.2 Evaluation of the Rate of Handovers

Given the Markov process characterization of the greedy-based association policy, one can evaluate the rate of handovers seen by a mobile user as stated in the following theorem:

Theorem 18. *Consider a mobile user moving along a straight line at constant velocity v in a Poisson cellular network with base station intensity, λ. Under the rectangle-based association policy with accessible base stations defined by $R_{(r_l,r_b)}(u)$, the rate of handovers seen by the mobile user, ρ is given by*

$$\rho = \frac{2\lambda r_b v}{2\lambda r_l r_b + q + 2\lambda r_b(1-2q)(-\mathcal{L}'_{\pi_c}(0))}, \tag{4.27}$$

where, $q = \pi_d(r_l) = \frac{1-\mathcal{L}_{\pi_c}(0)}{2}$ and $\mathcal{L}'_{\pi_c}(0)$ is given by (4.26).

Proof. Recall that $(T_i)_{i=1}^{\infty}$ denote the handover times and that at any time, T_n, X_n is the relative x-coordinate of the association base station. Then, we have that

$$T_1 = \begin{cases} \frac{\hat{X}_1 + r_l}{v} & \text{w.p. } 1-2q, \\ \frac{2r_l}{v} & \text{w.p } q, \\ G & \text{w.p. } q, \end{cases} \tag{4.28}$$

where, q is given in (4.23) and $G \sim \exp(2\lambda r_b v)$ is the time the rectangle $R_{(r_l, r_b)}(u)$ is empty of points. The rate of handovers is then given by:

$$\rho = \frac{1}{\mathbb{E}^0[T_1]}, \tag{4.29}$$

where $\mathbb{E}^0$ denotes the Palm expectation which is given by:

$$\begin{aligned} \mathbb{E}^0[T_1] &= (1-2q)\mathbb{E}^0\left[\frac{\hat{X}_1 + r_l}{v}\right] + \frac{2qr_l}{v} + \frac{q}{2\lambda r_b v} \\ &= \frac{r_l}{v} + \frac{q}{2\lambda r_b v} + \frac{1-2q}{v}(-\mathcal{L}'_{\pi_c}(0)), \end{aligned} \tag{4.30}$$

since, $\mathbb{E}^0[\hat{X}_1] = -\mathcal{L}'_{\pi_c}(0)$, which is given by (4.26).

□

Corollary 2. *Under the rectangle-based association policy, we have that*

$$\lim_{\lambda \to \infty} \rho = \frac{v}{2r_l}. \tag{4.31}$$

Proof. As the density of base stations, λ, goes to infinity, we have that $\lim_{\lambda\to\infty} \mathcal{L}_{\pi_c}(0) = 1$, and $\mathbb{E}[\hat{X}_1] = r_l$. Thus, we have that $q = 0$ and the rate of handovers is a constant $\frac{v}{2r_l}$. □

4.5.3 Evaluation of Effective Throughput

Given the steady state formulation of the Markov process, the expectation of the rate can be evaluated using the inversion formula [13] as:

$$\mathbb{S} = \mathbb{E}[S(0)] = \frac{\mathbb{E}^0\left[\int_0^{(T_1-\delta)^+} S(0)o\ \theta_t dt\right] - c}{\mathbb{E}^0[T_1]}, \tag{4.32}$$

where, θ_t is the shift operator [13]. For mathematical simplicity, we consider that the user is served at all times and we take the effect of handovers into account by introducing a cost term as follows:

$$\hat{\mathbb{S}} = \frac{\mathbb{E}^0\left[\int_0^{T_1} S(0)o\ \theta_t dt\right]}{\mathbb{E}^0[T_1]}(1-\rho\delta) - \rho c, \tag{4.33}$$

where $\rho = \frac{1}{\mathbb{E}^0[T_1]}$ is the rate of handovers and c is the fixed penalty for each handover. Note that the handover cost $\delta\rho$ is a unit-less quantity used to quantify the fraction of interruption time along the user trajectory.

First, let us consider the distance process

$$D(t) = \sqrt{(vt - X_n)^2 + Y_n^2}, \quad \text{for } T_n \leq t < T_{n+1}, \tag{4.34}$$

if it is connected and ∞ otherwise. Let the relative position of the association base station conditioned on the availability of an accessible base station be denoted by $(\hat{X}_1, \hat{Y}_1)$. Given the distribution of random variable T_1 as in (4.28), and $q = \pi_d(\infty)$, the expectation is evaluated as follows:

$$\begin{aligned}\mathbb{E}^0\left[\int_0^{T_1} S(0)o\ \theta_t dt\right] =& q\mathbb{E}^0\left[\int_0^{2r_l/v} h(r_l, \hat{Y}_1, t)dt\right] \\ &+(1-2q)\mathbb{E}^0\left[\int_0^{(\hat{X}_1+r_l)/v} g(\hat{X}_1, \hat{Y}_1, t)dt\right],\end{aligned} \tag{4.35}$$

where

$$g(\hat{X}_1, \hat{Y}_1, t) = \omega \ln \left(1 + \frac{\xi \times l(\sqrt{(vt - \hat{X}_1)^2 + \hat{Y}_1^2})}{\sigma^2} \right),$$

and

$$h(r_l, \hat{Y}_1, t) = \omega \ln \left(1 + \frac{\xi \times l(\sqrt{(vt - r_l)^2 + \hat{Y}_1^2})}{\sigma^2} \right).$$

Special Case. Let us evaluate the effective throughput for the bounded path loss function

$$l_2(x) = \exp(-\kappa x^2). \tag{4.36}$$

Approximating $\ln(1 + x) \sim \ln(x)$, for large x, we have that:

$$\begin{aligned} g(\hat{X}_1, \hat{Y}_1, t) &= \omega \left(\ln(\xi/\sigma^2) - k \left[(vt - \hat{X}_1)^2 + \hat{Y}_1^2 \right] \right), \\ h(r_l, \hat{Y}_1, t) &= \omega \left(\ln(\xi/\sigma^2) - k \left[(vt - r_l)^2 + \hat{Y}_1^2 \right] \right). \end{aligned} \tag{4.37}$$

Let $m = \ln(\xi/\sigma^2)$. Now, given $\hat{Y}_1 \sim U[-r_b, r_b]$, one can evaluate the expectation in (4.35), which gives:

$$\begin{aligned} \mathbb{E}[S(0)] =& \frac{q\omega}{\mathbb{E}^0[T_1]} \left[\frac{6r_l m - 2\kappa r_l^3 - 2r_l r_b^2 \kappa}{3v} \right] + \\ & \frac{(1-2q)\omega}{\mathbb{E}^0[T_1]} \left[\frac{3m\mathbb{E}^0[X_1] - \kappa \mathbb{E}^0[X_1^3] + 3mr_l - \kappa r_l^3 - \kappa r_l r_b^2}{3v} \right], \end{aligned} \tag{4.38}$$

where $q = \pi_d(\infty)$ is given in (4.23), and $\mathbb{E}^0[T_1]$ is given in (4.30).

4.5.4 Optimal Farthest Greedy Policy

The problem of finding the optimal policy is now posed as a parametric optimization problem as follows:

Problem 3. The problem of finding an optimal sequence of association base stations and handover times under rectangle-based farthest greedy association policy consists in finding the optimal parameters (r_l, r_b) of the region $R_{(r_l,r_b)}(u)$ that maximize the effective throughput $\mathcal{S}((r_l, r_b))$.

$$\mathcal{S}^* = \max_{r_l,r_b}\{\mathcal{S}((r_l, r_b))\}. \tag{4.39}$$

Associating with a base station at a greater distance decreases the handovers but may result in low rate. Thus, we focus on optimizing the stationary effective throughput with respect to the constraint on the greediness, i.e., the size of the association region.

If we consider a rectangle of equal side lengths, i.e, a square association region of side length $r = r_l = r_b$, we have that in the limit r going to zero and r going to ∞, the effective throughput in (4.35), goes to zero. Thus, a local or global optimum value must exist for the effective throughput. Note that it is challenging to give a closed form expression for the optimal parameters r_l^*, r_b^* which maximizes effective throughput, thus we evaluate the optimal parameter values with the help of simulations in Section VI.

4.6 Numerical Results and Simulations

In this section, we study the effectiveness of our proposed association policies with the help of simulation. We compare their performance and assess their improvements with respect to the closest base station association policy. We then study the performance sensitivity of these policies to the velocity, cost of handovers and handover delay. Unless otherwise stated, we consider a realization of a Poisson point process of intensity λ and the parameters of Table III. In particular, we will consider the following questions:

- What is the performance and complexity tradeoff for the optimal forward association policies with respect to the size of the association region $A_\gamma(u)$?
- What is the trade-off between the farthest distance and the height of the association region, i.e., the length and breadth in a rectangular region with respect to performance under greedy policies?
- What are the performance improvements of the optimal forward policy over the optimal greedy policy under various environments?
- When do the performance metrics of both policies have low discrepancy?

4.6.1 Optimal Forward Association Policy

In this section, we will consider an association region $A_\gamma(u)$ with the tagged user traveling a distance of 10Km to destination $(10, 0)$ on the positive

Parameter	Value
Power ξ	1 W
Bandwidth ω	30 MHz
Velocity v	20 m/s
Handover Delay δ	20 ms
Handover Cost c	100 Kbits

Table 4.3: Table of simulation parameters.

x-axis. We build the directed acyclic graph as in Definition 3 and solve the problem of the maximum weight path using Dijkstra's algorithm.

Problem 2 consists in finding the association sequence $\mathbf{p}^* \in \mathcal{P}_f$ that results in a maximum reward, i.e., the total volume of data delivered for a fixed association region $A_\gamma(u)$. Note that increasing the size of the association region leads to more feasible association sequences, i.e., more nodes in the graph and hence higher complexity. We notice that beyond a certain size of the association region, increasing the region does not improve the optimization. Thus for our simulation study, we consider one fixed size, $r_l = 400$m for the square region, the value of which depend on the decay of the path loss function considered.

4.6.2 Farthest Greedy Association Policy

For rectangular-based farthest greedy policies, we have a closed form expression for the effective throughput and the rate of handovers. In this subsection, we first compare the numerical results with the simulation. We then consider the dependence of the optimal parameters on the intensity of

base stations λ.

4.6.2.1 Evaluation of the Rate of Handovers

We numerically evaluated the closed form expression for the mean rate of handovers as given in (4.29), for the intensity of base stations, $\lambda = 40$ per Km^2. We consider a square association region, i.e., $r_l = r_b$ and compare the numerical results with that of the simulation. The simulation results agree with the numerical evaluation for various values of the side of the square as illustrated in Figure 4.9.

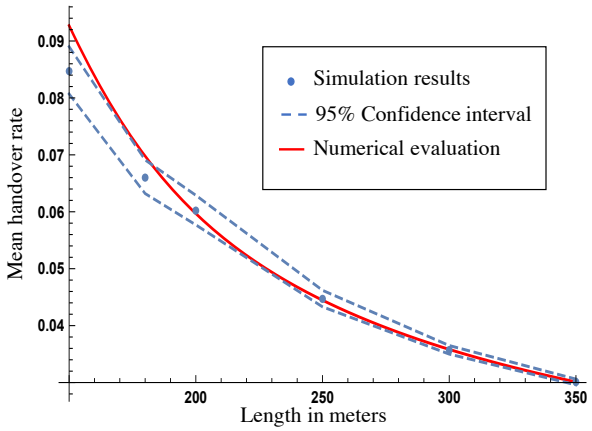


Figure 4.9: Comparison of the numerical results with simulation for a various lengths of the square region and $\lambda = 40$ per Km^2.

4.6.2.2 Evaluation of Effective Throughput

In this subsection, we focus on the tradeoff between the length and breadth of the region with respect to performance for differnet (1) intensities

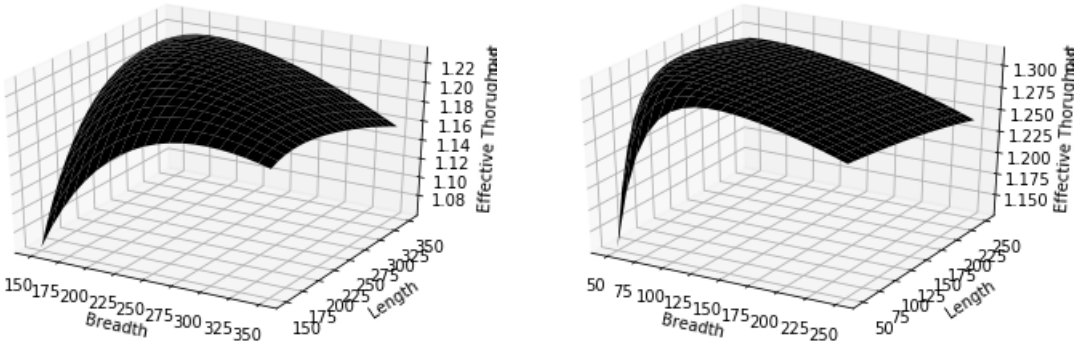
of base stations, and (2) path loss functions. While the cost of handover depends on how far in the direction of motion the user can associate with, i.e., the length of the region, the rate is seen by the user depends on how close the base station is to the trajectory, i.e., the breadth of the region.

Intensity of Base Stations. We numerically evaluate the effective throughput for various values of the parameters r_l, r_b for two intensities of base stations $\lambda = 20$ per Km^2 and $\lambda = 200$ per Km^2. Figure 4.10 illustrates the surface plots of the effective throughput for the path loss function given in (4.36).

We have two insights:(1) When increasing λ, the probability of finding a base station at the forward edge of the rectangular region increases, thus the optimal size is smaller for higher densities of base stations, and (2) Given that a base station that is closer to the tagged user's trajectory leads to a higher rate, the shape of the optimal regions is an elongated rectangle, i.e., the length is higher than the breadth.

4.6.2.3 Parametric Optimization under constraint

We now consider the parametric optimization problem under additional constraint that the rectangular handover region has a fixed area. There is a certain amount of complexity involved in gaining the knowledge of the base station locations. Thus we limit the complexity by considering a fixed area of the rectangular association region. In this setting, we evaluate the effective throughput with the help of simulations for various intensities of base station and an area of $4r_l r_b = 10^4\text{m}^2$, as illustrated in Figure 4.11, under the area



(a) (b)

Figure 4.10: Surface plots of effective throughput for 2 intensities of base stations. (a) For $\lambda = 20$, the optimal parameters $(r_l^*, r_b^*) = (286.8, 171)$m and (b) for $\lambda = 200$, the optimal parameters $(r_l^*, r_b^*) = (176.3, 113.1)$m.

constraint. The optimal rectangular shape is narrower for a higher intensity of base stations as before.

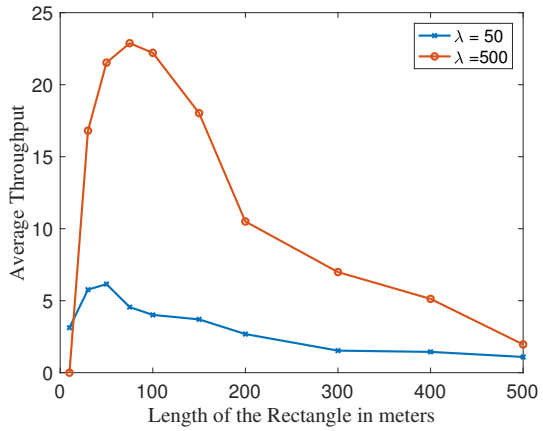


Figure 4.11: Comparing the effective throughput for various intensities of base stations under the constraints of fixed area.

4.6.3 Comparison of Performance of Optimal Forward Association and Optimal Farthest Greedy Policy

In this subsection, we compare the optimal farthest greedy policy of Section V with the dynamic programming based policy of Section IV by considering a common realization of a Poisson point process on which they are jointly applied. We consider rectangular association regions $R_{(r_l,r_b)}(u)$ with the optimal parameters for the greedy policies that are evaluated as in the previous section. We consider the total volume of data delivered as the common performance metric. To account for the variability of the Poisson point process, we average the total volume of data delivered for many realizations of the point process.

The global picture is that the optimal forward association policy has roughly the same performance as the greedy policy under the considered parameters. We evaluated the performance for intensity of base stations $\lambda = 20, 80, 200$ per Km^2. The percentage of increase in the total volume of data delivered by the dynamic programming approach compared to the greedy policy is in the range of $1 - 5\%$. The improvement tends to 0 when increasing (1) the handover delay, (2) the cost of each handover, or (3) the velocity of the user, as illustrated in Figures 4.12 - 4.14.

Further, we compare the performance of our policies with the traditional closest base station policy. We have that the improvement is larger for the high cost of handovers and handover delays in dense networks as illustrated in Figures 4.13 - 4.14.

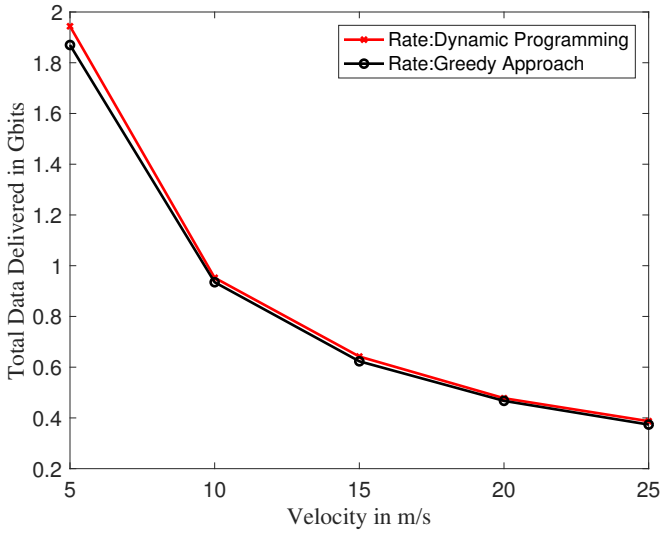


Figure 4.12: Total data delivered vs velocity ($\lambda = 80$ per Km^2).

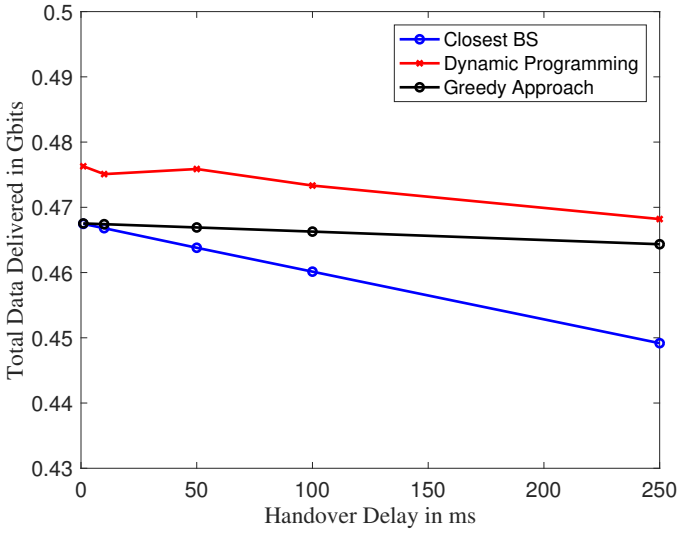


Figure 4.13: Total data delivered vs handover delay ($\lambda = 80$ per Km^2).

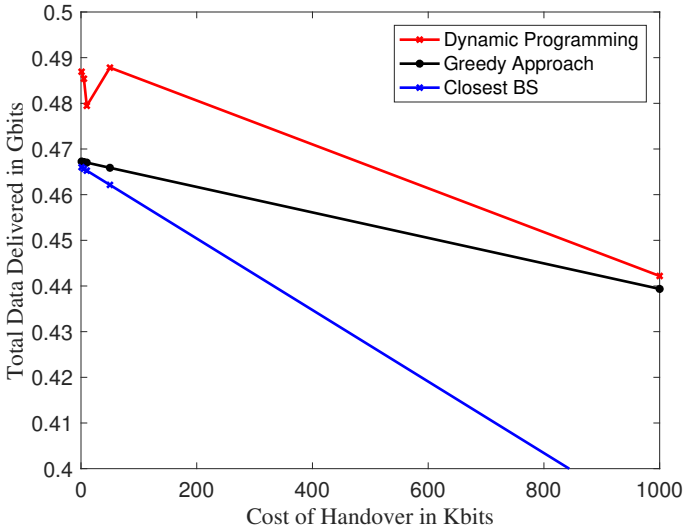


Figure 4.14: Total data delivered vs handover cost ($\lambda = 80$ per Km^2).

4.7 Conclusion

In ultra dense networks, the increased intensity of handovers has a negative impact on mobile users' performance. In this paper, we proposed strategies for association that achieve an appropriate throughput-handover trade-off. The first strategy takes the user's trajectory and the locations of all base stations into account, and determines an optimal set of association base stations and handover times. This is achieved by finding the maximum weight path in a directed acyclic graph.

Given only limited knowledge of the base station locations and user trajectory, the second strategy is a direction dependent greedy association policy that results in a reduced number of handovers. We give a closed form

expression for the stationary effective throughput seen by the user under this policy as a result of a Markovian formulation. This is in turn used to derive the optimal farthest greedy association policy.

We validate the effectiveness of the optimized greedy policy by comparing it with dynamic programming in various scenarios. The percentage gains achieved by the dynamic programming policy is within $1-5\%$ from the greedy policies for the considered parameters.

4.8 Appendix

4.8.1 Disc-based greedy policies

Under the the disc-based association policy, where the region defining the set of accessible base stations as given in (4.1), is a disc $D_r(u)$ of radius r, the state space of the Markov process $(\mathbf{Z}_n, n \in \mathbb{N})$ is given by:

$$\mathcal{Z} = \{[-r, r] \times [-r, r], \infty\}. \tag{4.40}$$

The state space $\mathcal{Z}$ is a mixture of continuous and discrete states, which implies that the transition probability kernel and the stationary distribution over the state space consists of a continuous part and a discrete part.

Suppose the previous state, $\mathbf{Z}_n = \mathbf{z} = (w, h)$, $w \in [-r, r]$, $h \in [-r, r]$ then the disc $D_r(u(T_n))$ around the mobile has two regions, R_1 where there must be no base stations and R_2 the remaining region of the square as illustrated in Figure 4.15. We evaluate the transition probabilities for $\mathbf{Z}_{n+1}$, with $Q(x, y) = [-r, x] \times [-r, y]$.

First, let us consider the discrete transition probability kernel, $\mathbb{P}(\mathbf{z}, \infty)$ which is given by the probability that region, $D_r(u(T_{n+1})) \setminus D_r(u(T_n))$, i.e., R_2 is empty of points:

$$\mathbb{P}(\mathbf{z}, \infty) = e^{-\lambda(\pi r^2 - a)}, \tag{4.41}$$

where a is the area of the region R_1, given by:

$$a = 2r^2 \cos^{-1}(\frac{d}{2r}) - d/2\sqrt{4r^2 - d^2}, \tag{4.42}$$

where $d = w + \sqrt{r^2 - h^2}$.

Given that the disc is empty of base stations, the first base station to enter will be at the edge of the disc. The y-coordinate of the base station is uniform between $[-r, r]$, thus $\mathbb{P}(\infty, Q(x, y))$ is given by:

$$\begin{aligned}\mathbb{P}(\infty, Q(x, y)) &= \mathbb{P}(X_{n+1} \leq x | Y_{n+1} = y, \mathbf{Z}_n = \infty) \times \mathbb{P}(Y_{n+1} \leq y | \mathbf{Z}_n = \infty), \\ &= \delta(x - \sqrt{r^2 - y^2}) \frac{y + r}{2r}.\end{aligned} \tag{4.43}$$

Given $Q((x, y)) = \{[-r, x] \times [-r, y]\}$, the region R_2 is further split into two regions R_3 and R_4. Then, the transition probability $\mathbb{P}(X_{n+1} =\leq x | \mathbf{Z}_n = (w, h))$ is given by the product of probability that the region R_3 has at least one point and the region R_4 has no points, see Fig 4.15:

$$\mathbb{P}(X_{n+1} = x | X_n = w, Y_n = h) = e^{-\lambda b}\left(1 - e^{-\lambda(\pi r^2 - a - b)}\right), \tag{4.44}$$

where a is given in (4.42) and b is the area of the regionR_4 given as,

$$b = r^2/2(2\cos^{-1}(x/r) - r^2/2\sin(2\cos^{-1}(x/r)). \tag{4.45}$$

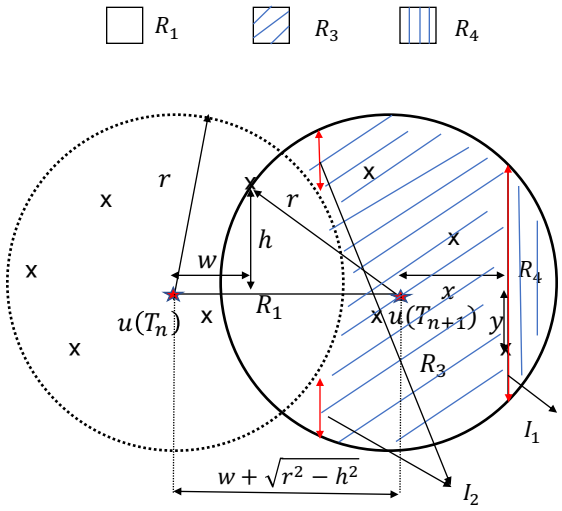


Figure 4.15: Two regions R_1 and $R_2 = R_3 \cup R_4$ of the discs.

Given that $X_n = x$, the distribution of Y_n is uniform random variable on the interval I_1 for $x \in [r-(w+\sqrt{r^2-h^2}), r]$ and for $x < r-(w+\sqrt{r^2-h^2})$, it is uniform on the interval I_2 as illustrated in Figure 4.15. Thus, the transition probability is given by:

$$\mathbb{P}(Y_{n+1} \leq y | X_n = x, \mathbf{Z}_n = (w,h)) =$$

$$\begin{cases} \frac{y+\sqrt{r^2-x^2}}{2\sqrt{r^2-x^2}} & x \in [r-(w+\sqrt{r^2-h^2}), r] \\ \frac{2+(y-\sqrt{r^2-(w+\sqrt{r^2-h^2}-x)^2})}{2(\sqrt{r^2-(w+\sqrt{r^2-h^2}-x)^2}-\sqrt{r^2-x^2})} & x < r-(w+\sqrt{r^2-h^2}) \end{cases} \tag{4.46}$$

The transition kernel density $f(\mathbf{z}_n, Q(x,y)))$ is given by the product of differention of the transition probabilities in (4.44) and (4.46) with respect to x anf y respectively.

4.8.2 Proof of Theorem 17

Let us consider the stationary distribution of this process in terms of π_d and π_c associated with the discrete and the continuous parts. Thus, $\pi_c(x)$ represents the continuous part of the stationary distribution of the Markov process for $x \in [-r_l, r_l)$, $\pi_d(r_l)$ and $\pi_d(\infty)$ are the discrete probability masses associated with the states r_l and ∞.We have the following equations:

$$\begin{aligned}
\pi_c(x) &= 2\lambda r_b e^{-2\lambda r_b(r_l - x)} \left(\int_{-x}^{r_l^-} \pi_c(w)dw + \pi_d(r_l) \right), \\
\pi_d(\infty) &= e^{-2\lambda r_b r_l} \int_{-r_l}^{r_l^-} \pi_c(w) e^{-2\lambda r_b w} dw + \pi_d(r_l) e^{-4\lambda r_b r_l}, \\
\pi_d(r_l) &= \pi_d(\infty),
\end{aligned} \tag{4.47}$$

with $\pi_d(r_l) + \pi_d(\infty) + \int_{-r_l}^{r_l^-} \pi_c(w)dw = 1$. We use Laplace transform techniques to solve the integral equations. First, we multiply on both sides of the first equation in (4.47) with e^{-sx} and integrate from $-r_l$ to r_l, we get

$$\begin{aligned}
\int_{-r_l}^{r_l} \pi_c(x) e^{-sx} dx = \frac{2\lambda r_b}{(2\lambda r_b - s)} \Bigg[& e^{-sr_l} - q e^{-4\lambda r_b r_l + sr_l} + \\
& e^{-sr_l} \int_{-r_l}^{r_l} \pi_c(z)dz - e^{-2\lambda r_b r_l} \int_{-r_l}^{r_l} \pi_c(z) e^{-(2\lambda r_b - s)z} dz \Bigg]
\end{aligned} \tag{4.48}$$

Let $\mathcal{L}_{\pi_c}(s) = \int_{-r_l}^{r_l} \pi_c(x) e^{-sx} dx$, then we have

$$\mathcal{L}_{\pi_c}(s) = \frac{2\lambda r_b}{(2\lambda r_b - s)} \left[e^{-sr} \mathcal{L}_{\pi_c}(0) - e^{-2\lambda r_b r_l} \mathcal{L}_{\pi_c}(2\lambda r_b - s) + q e^{-sr_l} - q e^{-4\lambda r_b r_l + sr_l} \right], \tag{4.49}$$

where, $q = \pi_d(r) = \frac{1 - \mathcal{L}_{\pi_c}(0)}{2}$.

Let $\alpha(s)$ and $\beta(s)$ be two functions such that

$$\begin{aligned}\mathcal{L}_{\pi_c}(s) &= \alpha(s) + \beta(s)\mathcal{L}_{\pi_c}(2\lambda r_b - s) \\ &= \alpha(s) + \beta(s)[\alpha(2\lambda r_b - s) + \beta(2\lambda r - s)\mathcal{L}_{\pi_c}(s)] \\ &= \frac{\alpha(s) + \beta(s)\alpha(2\lambda r_b - s)}{1 - \beta(s)\beta(2\lambda r_b - s)}.\end{aligned} \tag{4.50}$$

From (4.49) we have:

$$\begin{aligned}\alpha(s) &= \frac{2\lambda r_b}{(2\lambda r_b - s)}\left[e^{-sr_l}\mathcal{L}_{\pi_c}(0) + qe^{-sr_l} - qe^{-4\lambda r_b r_l + sr_l}\right], \\ \beta(s) &= -\frac{2\lambda r_b e^{-2\lambda r_b r_l}}{(2\lambda r_b - s)}.\end{aligned} \tag{4.51}$$

Since the range of the random variable with probability distribution π_c is finite, the Laplace transform cannot be degenerate. Therefore, the zeros of the denominator in (4.50), should match the zeros of the numerator. Thus, the zeros a_1, a_2 of the denominator are given by solving $1 = \beta(s)\beta(2\lambda r_b - s)$ which gives the following quadratic equation

$$s^2 - 2\lambda r_b s + 4\lambda^2 r_b^2 e^{-4\lambda r_b r_l} = 0. \tag{4.52}$$

Thus, the roots of the denominator are $a_1, a_2 = \lambda r_b \pm \lambda r_b\sqrt{1 - 4e^{-4\lambda r_b r_l}}$. Note, that the roots are imaginary for $4\lambda r_b r_l < 1.38$. Thus, the mean number of base stations($\lambda 4 r_b r_l$) in the rectangular region of side lengths $2r_l, 2r_b$ considered has to be greater than 1.38 for real roots.

Now, let us denote the numerator in (4.50) as $f(s)$, then we have that $f(a_1) = f(a_2) = 0$, using which we will solve for $\mathcal{L}(0)$. The function $f(s)$ is

given as:

$$f(s) = \left[2\lambda r_b s e^{-sr_l} - 4\lambda^2 r_b^2 e^{-4\lambda r_b r_l + sr_l}\right]\left(\frac{1+\mathcal{L}_{\pi_c}(0)}{2}\right) + \left[4\lambda^2 r_b^2 e^{-4\lambda r_b r_l - sr_l} - 2\lambda s r_b e^{-4\lambda r_b r_l + sr_l}\right]\left(\frac{1-\mathcal{L}_{\pi_c}(0)}{2}\right). \tag{4.53}$$

Then, $f(\lambda r_b + \lambda r_b\sqrt{1-4e^{-4\lambda r_b r_l}}) = 0$, yields (4.54).

$$\mathcal{L}_{\pi_c}(0) = \frac{2\lambda r_b\left(1-e^{2a_1 r_l}\right) + a_1\left(e^{4\lambda r_l r_b} - e^{2a_1 r_l}\right)}{2\lambda r_b\left(1+e^{2a_1 r_l}\right) - a_1\left(e^{4\lambda r_l r_b} + e^{2a_1 r_l}\right)}, \tag{4.54}$$

where, $a_1 = \lambda r + \lambda r\sqrt{1-4e^{-4\lambda r^2}}$.

Similarly for the root $a_2 = \lambda r - \lambda r\sqrt{1-4e^{-4\lambda r^2}}$, we get:

$$\mathcal{L}_{\pi_c}(0) = \frac{2\lambda r_b\left(1-e^{2a_2 r_l}\right) + a_2\left(e^{4\lambda r_l r_b} - e^{2a_2 r_l}\right)}{2\lambda r_b\left(1+e^{2a_2 r_l}\right) - a_2\left(e^{4\lambda r_l r_b} + e^{2a_2 r_l}\right)}, \tag{4.55}$$

Let us prove that both the above functions are same. Since a_1 and a_2 are roots of the quadratic equation given in (4.52), we have that:

$$\begin{aligned} a_1 + a_2 &= 2\lambda r_b, \\ a_1 a_2 &= 4\lambda^2 r_b^2 e^{-4\lambda r_b r_l}. \end{aligned} \tag{4.56}$$

Let us substitute $2\lambda r_b$ in (4.54) with $\frac{a_1 a_2 e^{4\lambda r_b r_l}}{2\lambda r_b}$:

$$\begin{aligned}\mathcal{L}_{\pi_c}(0) &= \frac{a_2 e^{4\lambda r_b r_l}\Big(1 - e^{2a_1 r_l}\Big) + 2\lambda r_b\Big(e^{4\lambda r_l r_b} - e^{2a_1 r_l}\Big)}{a_2 e^{4\lambda r_b r_l}\Big(1 + e^{2a_1 r_l}\Big) - 2\lambda r_b\Big(e^{4\lambda r_l r_b} + e^{2a_1 r_l}\Big)}, \\ &= \frac{a_2\Big(1 - e^{4\lambda r_l r_b} e^{-2a_2 r_l}\Big) + 2\lambda r_b\Big(1 - e^{-2a_2 r_l}\Big)}{a_2\Big(1 + e^{4\lambda r_l r_b} e^{-2a_2 r_l}\Big) - 2\lambda r_b\Big(1 + e^{-2a_2 r_l}\Big)},\end{aligned} \tag{4.57}$$

where the last equality is obtained by substituting $a_1 = 2\lambda r_b - a_2$. Now, we get (4.55) by multiplying the numerator and the denominator of the above equation by $-e^{2a_2 r_l}$.

Chapter 5

Conclusion

In this thesis, we took several steps toward quantifying the performance metrics seen by a population of mobile users by characterizing the spatial stochastic fields associated with wireless networks. Specifically, in the first two chapters, we characterized temporal variations seen by a typical mobile user in various environments such as noise-limited and dense interference limited networks. This is a challenging problem and thus our focus in this thesis was limited to studying the impact of large scale properties of the system, e.g., the density of base stations, path loss functions, and partially addressing the impact of multipath-fading and shadowing via simulation.

We successfully characterized Gaussian limits of the interference field in dense wireless networks and studied the level sets and spatial correlations of the Shannon rate. Although we neglected fading, we observed significant variability in the Shannon rate in the case where the path loss functions have discontinuities possibly associated with environmental blockages when operating in high frequencies. We characterized such variability using Hölders exponent. We then focused on studying how such capacity variability could be smoothed out via buffering and/or conservative rate modulation techniques

required to tolerate/adapt to such variability. However, high variability leads to a significant reduction in the efficiency of wireless systems particularly when they support a significant volume of low-latency traffic.

In addition to enabling an evaluation of performance for mobile users, the characterization of spatial fields also leads to quantifying system-level performance metrics such as the required backhaul capacity at a gateway. In this thesis, we studied the impact that variability (second moments) in the rate and spatial correlations have on such system-level performance and backhaul costs.

In dense networks, mobile users may experience rapid handovers and associated overheads which may be reduced through intelligent handover decisions. In the last chapter, we proposed a class of greedy association policies for mobile users in a dense network. Mobile users can make proactive mobility driven handover decisions, in the case where offloading mobile users to macro-base stations leads to congestion due to a large volume.

In this thesis, we proposed new models and a new class of association policies leveraging stochastic geometry and the theory of Markov processes. The proposed policies perform better than the traditional closest base station policy, particularly in cases with high handover costs. Still, it is desirable for dense systems to focus on reducing handover costs, e.g., soft handovers since highly mobile users in dense networks pay a significant cost with frequent handovers. The proposed association policies focused on realizing the tradeoff between throughput and handover costs. In the future, it would be desirable

to extend this to include load balancing along with co-optimization of base station activation policies which aim to conserve energy.

Bibliography

[1] Cisco Visual Networking Index. Global mobile data traffic forecast update, 2014–2019. *White Paper, February*, 1, 2015.

[2] Wandera. Enterprise Mobile Data Report - q3. Oct 2013.

[3] Jeffrey G Andrews, Stefano Buzzi, Wan Choi, Stephen V Hanly, Angel Lozano, Anthony CK Soong, and Jianzhong Charlie Zhang. What will 5G be? *IEEE Journal on selected areas in communications*, 32(6):1065–1082, 2014.

[4] Fung Po Tso, Jin Teng, Weijia Jia, and Dong Xuan. Mobility: a double-edged sword for HSPA networks: a large-scale test on Hong Kong mobile HSPA networks. In *Proceedings of the eleventh ACM international symposium on Mobile ad hoc networking and computing*, pages 81–90. ACM, 2010.

[5] Petteri Lunden, Jussi Aijanen, Kari Aho, and Tapani Ristaniemi. Performance of VoIP over HSDPA in mobility scenarios. In *Vehicular Technology Conference, 2008. VTC Spring 2008. IEEE*, pages 2046–2050. IEEE, 2008.

[6] François Baccelli and Bartlomiej Blaszczyszyn. Spatial modeling of wireless communications: a stochastic geometry approach. *Foundations and*

Trends in Networking, NOW Publishers, 2009.

[7] "D.J. Aldous". *"Probability Approximations via the Poisson Clumping Heuristic"*. Springer-Verlag, 1989.

[8] Julian Keilson. *Markov chain model-rarity and exponentiality*, volume 28. Springer Science & Business Media, 2012.

[9] Andrea Goldsmith. *Wireless communications*. Cambridge university press, 2005.

[10] Martin J Feuerstein, Kenneth L Blackard, Theodore S Rappaport, Scott Y Seidel, and Howard H Xia. Path loss, delay spread, and outage models as functions of antenna height for microcellular system design. *IEEE Transactions on Vehicular Technology*, 43(3):487–498, 1994.

[11] Amitava Ghosh, Timothy A Thomas, Mark C Cudak, Rapeepat Ratasuk, Prakash Moorut, Frederick W Vook, Theodore S Rappaport, George R MacCartney, Shu Sun, and Shuai Nie. Millimeter-wave enhanced local area systems: A high-data-rate approach for future wireless networks. *IEEE Journal on Selected Areas in Communications*, 32(6):1152–1163, 2014.

[12] Yu K Belyaev. Continuity and hölder's conditions for sample functions of stationary gaussian processes. In *Proc. 4th Berkeley Symp. Math. Statist. and Prob*, volume 2, pages 22–33, 1961.

[13] Francois Baccelli and Pierre Brémaud. *Elements of queueing theory: Palm Martingale calculus and stochastic recurrences*, volume 26. Springer Science and Business Media, 2013.

[14] Francois Baccelli and B. Blaszczyszyn. *Spatial Modeling of Wireless Communications – A Stochastic Geometry Approach*. NOW Publishers, 2009.

[15] Thomas Karagiannis, Jean-Yves Le Boudec, and Milan Vojnović. Power law and exponential decay of intercontact times between mobile devices. *Mobile Computing, IEEE Transactions on*, 2010.

[16] Vania Conan, Jérémie Leguay, and Timur Friedman. Characterizing pairwise inter-contact patterns in delay tolerant networks. In *Proceedings of the 1st international conference on Autonomic computing and communication systems*. ICST, 2007.

[17] Han Cai and Do Young Eun. Toward stochastic anatomy of inter-meeting time distribution under general mobility models. In *Proceedings of the 9th ACM international symposium on Mobile ad hoc networking and computing*. ACM, 2008.

[18] Shubhadip Mitra, Sayan Ranu, Vinay Kolar, Aditya Telang, Arnab Bhattacharya, Ravi Kokku, and Sriram Raghavan. Trajectory aware macrocell planning for mobile users. *arXiv preprint arXiv:1501.02918*, 2015.

[19] Han Deng and I-Hong Hou. Online scheduling for delayed mobile offloading. In *Proc. IEEE INFOCOM, 2015*.

[20] Zheng Lu and G. de Veciana. Optimizing stored video delivery for mobile networks: The value of knowing the future. In *Proc. IEEE INFOCOM, 2013*.

[21] Armand M Makowski. On a random sum formula for the busy period of the M/G/Infinity queue with applications. Technical report, DTIC Document, 2001.

[22] Sung Nok Chiu, Dietrich Stoyan, Wilfrid S Kendall, and Joseph Mecke. *Stochastic geometry and its applications*. John Wiley & Sons, 2013.

[23] Atsuyuki Okabe, Barry Boots, Kokichi Sugihara, and Sung Nok Chiu. *Spatial tessellations: concepts and applications of Voronoi diagrams*. John Wiley & Sons, 2009.

[24] Francois Baccelli and P. Brémaud. *Palm probabilities and Stationary Queues*. Springer Verlag, March 1987.

[25] Jesper Møller. Random johnson-mehl tessellations. *Advances in applied probability*, 1992.

[26] Eric W. Weisstein. Kolmogorov-Smirnov test. *MathWorld–A Wolfram Web Resource*.

[27] David B Johnson and David A Maltz. Dynamic source routing in ad hoc wireless networks. *Mobile computing*, pages 153–181, 1996.

[28] Varun Gupta and Peter G Harrison. Fluid level in a reservoir with an on-off source. *ACM SIGMETRICS Performance Evaluation Review*, 2008.

[29] John Frank Charles Kingman. *Poisson processes*. Wiley Online Library, 1993.

[30] Frédéric Morlot. A population model based on a poisson line tessellation. In *WiOpt'12: Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks*. IEEE, 2012.

[31] François Baccelli, Maurice Klein, Marc Lebourges, and Sergei Zuyev. Stochastic geometry and architecture of communication networks. *Telecommunication Systems*, 1997.

[32] Martin Haenggi, Jeffrey G Andrews, François Baccelli, Olivier Dousse, and Massimo Franceschetti. Stochastic geometry and random graphs for the analysis and design of wireless networks. *IEEE Journal on Selected Areas in Communications*, 27(7), 2009.

[33] Jeffrey G Andrews, François Baccelli, and Radha Krishna Ganti. A tractable approach to coverage and rate in cellular networks. *IEEE Transactions on Communications*, 59(11):3122–3134, 2011.

[34] Van Minh Nguyen and Marios Kountouris. Performance limits of network densification. *IEEE Journal on Selected Areas in Communications*, 35(6):1294–1308, 2017.

[35] Xinchen Zhang and Jeffrey G Andrews. Downlink cellular network analysis with multi-slope path loss models. *IEEE Transactions on Communications*, 63(5):1881–1894, 2015.

[36] Chung Shue Chen, Van Minh Nguyen, and Laurent Thomas. On small cell network deployment: A comparative study of random and grid topologies. In *Vehicular Technology Conference (VTC Fall), 2012 IEEE*, pages 1–5. IEEE, 2012.

[37] Jesús Arnau, Italo Atzeni, and Marios Kountouris. Impact of LOS/NLOS propagation and path loss in ultra-dense cellular networks. In *Communications (ICC), 2016 IEEE International Conference on*, pages 1–6. IEEE, 2016.

[38] Ming Ding, Peng Wang, David López-Pérez, Guoqiang Mao, and Zihuai Lin. Performance impact of LoS and NLoS transmissions in dense cellular networks. *IEEE Transactions on Wireless Communications*, 15(3):2365–2380, 2016.

[39] François Baccelli and Xinchen Zhang. A correlated shadowing model for urban wireless networks. In *Computer Communications (INFOCOM), 2015 IEEE Conference on*, pages 801–809. IEEE, 2015.

[40] Radha Krishna Ganti and Martin Haenggi. Spatial and temporal correlation of the interference in ALOHA ad hoc networks. *IEEE Communications Letters*, 13(9), 2009.

[41] Udo Schilcher, Christian Bettstetter, and Günther Brandner. Temporal correlation of interference in wireless networks with rayleigh block fading. *IEEE Transactions on Mobile Computing*, 11(12):2109–2120, 2012.

[42] François Baccelli and Anup Biswas. On scaling limits of power law shot-noise fields. *Stochastic Models*, 31(2):187–207, 2015.

[43] Lothar Heinrich and Volker Schmidt. Normal convergence of multidimensional shot noise and rates of this convergence. *Advances in Applied Probability*, 17(04):709–730, 1985.

[44] At Papoulis. High density shot noise and gaussianity. *Journal of Applied Probability*, 8(1):118–127, 1971.

[45] Jean-Marc Azaïs and Mario Wschebor. *Level sets and extrema of random processes and fields*. John Wiley & Sons, 2009.

[46] Israel Bar-David and A Nemirovsky. Level crossings of nondifferentiable shot processes. *IEEE Transactions on Information Theory*, 18(1):27–34, 1972.

[47] Hermine Biermé, Agnès Desolneux, et al. Crossings of smooth shot noise processes. *The Annals of Applied Probability*, 22(6):2240–2281, 2012.

[48] Hermine Biermé and Agnes Desolneux. A fourier approach for the level crossings of shot noise processes with jumps. *Journal of Applied Probability*, 49(1):100–113, 2012.

[49] Harald Cramér and M Ross Leadbetter. *Stationary and related stochastic processes: Sample function properties and their applications*. Courier Corporation, 2013.

[50] Xiaohu Ge, Hui Cheng, Mohsen Guizani, and Tao Han. 5G wireless backhaul networks: Challenges and research advances. *IEEE Network*, 28(6):6–11, 2014.

[51] Sooyoung Hur, Taejoon Kim, David J Love, James V Krogmeier, Timothy A Thomas, Amitava Ghosh, et al. Millimeter wave beamforming for wireless backhaul and access in small cell networks. *IEEE Trans. Communications*, 61(10):4391–4403, 2013.

[52] Cedric Dehos, Jose Luis González, Antonio De Domenico, Dimitri Ktenas, and Laurent Dussopt. Millimeter-wave access and backhauling: the solution to the exponential data traffic increase in 5G mobile communications systems? *IEEE Communications Magazine*, 52(9):88–95, 2014.

[53] Krishna Balachandran, Srinivas R Kadaba, and Sanjiv Nanda. Channel quality estimation and rate adaptation for cellular mobile radio. *IEEE Journal on Selected Areas in Communications*, 17(7):1244–1256, 1999.

[54] Qingwen Liu, Shengli Zhou, and Georgios B Giannakis. Cross-layer combining of adaptive modulation and coding with truncated ARQ over wireless links. *IEEE Transactions on wireless communications*, 3(5):1746–1755, 2004.

[55] B François and B Bartłomiej. Stochastic geometry and wireless networks. volume 1. theory, 2009.

[56] Peter J Bickel and Michael J Wichura. Convergence criteria for multiparameter stochastic processes and some applications. *The Annals of Mathematical Statistics*, pages 1656–1670, 1971.

[57] Ahmad AlAmmouri, Jeffrey G Andrews, and François Baccelli. SINR and throughput of dense cellular networks with stretched exponential path loss. *arXiv preprint arXiv:1703.08246*, 2017.

[58] Vinko Erceg, Saeed Ghassemzadeh, Maxwell Taylor, Dong Li, and Donald L Schilling. Urban/suburban out-of-sight propagation modeling. *IEEE Communications Magazine*, 30(6):56–61, 1992.

[59] Pranav Madadi, Francois Baccelli, and Gustavo de Veciana. On temporal variations in mobile user SNR with applications to perceived QoS. In *Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks (WiOpt), 2016 14th International Symposium on*, pages 1–8. IEEE, 2016.

[60] Donald L Iglehart. Limiting diffusion approximations for the many server queue and the repairman problem. *Journal of Applied Probability*,

2(02):429–441, 1965.

[61] AA Borovkov. On limit laws for service processes in multi-channel systems. *Siberian Mathematical Journal*, 8(5):746–763, 1967.

[62] Anne Estrade, Ileana Iribarren, and Marie Kratz. Chord-length distribution functions and rice formulae. Application to random media. *Extremes*, 15(3):333–352, 2012.

[63] Georg Lindgren. *Stationary stochastic processes: theory and applications*. CRC Press, 2012.

[64] Robert J Adler and Jonathan E Taylor. *Random fields and geometry*. Springer Science & Business Media, 2009.

[65] Neil Sinclair, David Harle, Ian A Glover, and Robert C Atkinson. A kernel methods approach to reducing handover occurrences within LTE. In *Wireless Conference (European Wireless), 2012 18th European*, pages 1–8. VDE, 2012.

[66] Peng-Yong Kong and Andrzej Sluzek. Average packet delay analysis for a mobile user in a two-tier heterogeneous cellular network. *IEEE Systems Journal*, 2015.

[67] Haijun Zhang, Wenmin Ma, Wei Li, Wei Zheng, Xiangming Wen, and Chunxiao Jiang. Signalling cost evaluation of handover management schemes in LTE-advanced femtocell. In *Vehicular Technology Conference (VTC Spring), 2011 IEEE 73rd*, pages 1–5. IEEE, 2011.

[68] Haijun Zhang, Chunxiao Jiang, Julian Cheng, and Victor CM Leung. Cooperative interference mitigation and handover management for heterogeneous cloud small cell networks. *IEEE Wireless Communications*, 22(3):92–99, 2015.

[69] Sennur Ulukus and Gregory P Pollini. Handover delay in cellular wireless systems. In *Communications, 1998. ICC 98. Conference Record. 1998 IEEE International Conference on*, volume 3, pages 1370–1374. IEEE, 1998.

[70] Balaji Rengarajan and Gustavo De Veciana. Practical adaptive user association policies for wireless systems with dynamic interference. *IEEE/ACM Transactions on Networking*, 19(6):1690–1703, 2011.

[71] Qiaoyang Ye, Beiyu Rong, Yudong Chen, Mazin Al-Shalash, Constantine Caramanis, and Jeffrey G Andrews. User association for load balancing in heterogeneous cellular networks. *IEEE Transactions on Wireless Communications*, 12(6):2706–2716, 2013.

[72] Suman Das, Harish Viswanathan, and Gee Rittenhouse. Dynamic load balancing through coordinated scheduling in packet data systems. In *INFOCOM 2003. Twenty-Second Annual Joint Conference of the IEEE Computer and Communications. IEEE Societies*, volume 1, pages 786–796. IEEE, 2003.

[73] Thomas Bonald, Sem Borst, and Alexandre Proutiere. Inter-cell scheduling in wireless data networks. In *Wireless Conference 2005-Next Genera-*

tion Wireless and Mobile Communications and Services (European Wireless), 11th European, pages 1–7. VDE, 2005.

[74] Aimin Sang, Xiaodong Wang, Mohammad Madihian, and Richard D Gitlin. Coordinated load balancing, handoff/cell-site selection, and scheduling in multi-cell packet data systems. *Wireless Networks*, 14(1):103–120, 2008.

[75] Tian Bu Li Erran Li Ramachandran Ramjee, T Bu, and LE Li. Generalized proportional fair scheduling in third generation wireless data networks. In *IEEE INFOCOM*, pages 1–12, 2006.

[76] Hongseok Kim, Gustavo de Veciana, Xiangying Yang, and Muthaiah Venkatachalam. Alpha-optimal user association and cell load balancing in wireless networks. In *INFOCOM, 2010 Proceedings IEEE*, pages 1–5. IEEE, 2010.

[77] Kyuho Son, Song Chong, and Gustavo De Veciana. Dynamic association for load balancing and interference avoidance in multi-cell networks. *IEEE Transactions on Wireless Communications*, 8(7), 2009.

[78] Kyuho Son, Hongseok Kim, Yung Yi, and Bhaskar Krishnamachari. Base station operation and user association mechanisms for energy-delay tradeoffs in green cellular networks. *IEEE journal on selected areas in communications*, 29(8):1525–1536, 2011.

[79] François Baccelli, Bartlomiej Blaszczyszyn, and Paul Muhlethaler. An Aloha protocol for multihop mobile wireless networks. *IEEE Transactions on Information Theory*, 52(2):421–436, 2006.

[80] Xingqin Lin, Radha Krishna Ganti, Philip J Fleming, and Jeffrey G Andrews. Towards understanding the fundamentals of mobility in cellular networks. *IEEE Transactions on Wireless Communications*, 12(4):1686–1698, 2013.

[81] Wei Bao and Ben Liang. Stochastic geometric analysis of user mobility in heterogeneous wireless networks. *IEEE Journal on Selected Areas in Communications*, 33(10):2212–2225, 2015.

[82] Rabe Arshad, Hesham ElSawy, Sameh Sorour, Tareq Y Al-Naffouri, and Mohamed-Slim Alouini. Handover management in dense cellular networks: A stochastic geometry approach. In *Communications (ICC), 2016 IEEE International Conference on*, pages 1–7. IEEE, 2016.

[83] Rabe Arshad, Hesham ElSawy, Sameh Sorour, Tareq Y Al-Naffouri, and Mohamed-Slim Alouini. Cooperative handover management in dense cellular networks. In *Global Communications Conference (GLOBECOM), 2016 IEEE*, pages 1–6. IEEE, 2016.

[84] Konstantinos Dimou, Min Wang, Yu Yang, Muhammmad Kazmi, Anna Larmo, Jonas Pettersson, Walter Muller, and Ylva Timner. Handover within 3GPP LTE: design principles and performance. In *Vehicular*

Technology Conference Fall (VTC 2009-Fall), 2009 IEEE 70th, pages 1–5. IEEE, 2009.

[85] Toktam Mahmoodi and Srini Seetharaman. On using a SDN-based control plane in 5G mobile networks. In *Wireless World Research Forum, 32nd Meeting*. Citeseer, 2014.

[86] Thomas H Cormen, Charles E Leiserson, Ronald L Rivest, and Clifford Stein. Introduction to algorithms mit press. *Cambridge, MA*, page 819, 2001.

[87] Ravindra K Ahuja, Thomas L Magnanti, and James B Orlin. Network flows. 1988.

[88] Gareth O Roberts, Jeffrey S Rosenthal, et al. General state space Markov chains and MCMC algorithms. *Probability Surveys*, 1:20–71, 2004.

Vita

Pranav Madadi has received his bachelor degree from Indian Institute of Technology, Hyderabad in 2014. He has been pursuing Ph.D. degree at the Electrical and Computer Engineering department of The University of Texas at Austin, working under the supervision of Prof. François Baccelli and Prof. Gustavo de Veciana. His research interest is in analyzing future wireless networks using the tools of stochastic geometry and information theory, with primary focus on analyzing the performance of a mobile users on the move.

E-mail: pranavm1903@gmail.com

This dissertation was typeset with LaTeX[†] by the author.

[†]LaTeX is a document preparation system developed by Leslie Lamport as a special version of Donald Knuth's TeX Program.